Mauritius Research Council

INNOVATION FOR TECHNOLOGY

# TOURISM DEMAND FOR THE LENGTH OF STAY IN MAURITIUS: SOCIO-ECONOMIC STATUS AND PSYCHOLOGICAL DEVELOPMENT

## Final Report

*July 2018*

## Mauritius Research Council

*Address:*

Level 6, Ebene Heights
34, Cybercity
Ebene

Telephone: (230) 465 1235
Fax: (230) 465 1239
e-mail: mrc@intnet.mu
Website: www.mrc.org.mu

**This report is based on work supported by the Mauritius Research Council under award number MRC/RUN-1606. Any opinions, findings, recommendations and conclusions expressed herein are the author's and do not necessarily reflect those of the Council.**

# Table of Contents

# ACKNOWLEDGEMENT

The authors acknowledge the efforts of all the staff who have contributed and helped in a way or the other in this study.

Special appreciation to all the field workers and interviewers at the SSR International Airport and other representatives who helped in identifying the passengers and facilitating the interview.

The group of researchers is also grateful to Mr Suddul (Department of Civil Aviation) and Mr. Rajesh Luchman (Airports of Terminal Operation) for their support during data collection. Deep and sincere gratitude is extended to all the tourists for sparing their time to participate actively in the interviews. They were found to be available and cooperative.

Last but not least, the group of researchers would like to thank the Mauritius Research Council for the trust and the opportunity given to us to carry out the study.

# EXECUTIVE SUMMARY

This report analyses the determinants of the length of stay of international tourists taking vacations in Mauritius, based on a questionnaire distributed in the departure lounge of the Sir Seewoosagur Ramgoolam International Airport. A survival model is adopted to measure the relationship between duration of stay and the different covariates. It was concluded that the length of stay at the destination is related to the various demographic characteristics, rating of trip attributes, sustainability practices and personality traits of the respondents. Thus the need for taking into consideration these factors for marketing purposes and eventually the sustainability of tourism industry in Mauritius is being advocated for.

Keywords: duration models, tourism, Mauritius.

# CHAPTER 1: INTRODUCTION

## 1.1 Research Problem

Tourism has become one of the most prominent economic trends for numerous countries. For several destinations, this trend will persist to rise and tourism, per se, will become the most dynamic and fastest growing sector of the economy. Owing to this vital contribution, tourism planning is essential for the whole development process. The concept of planning explores some form of decision-making involving the future as tourism demand is the foundation whereby all tourism-related business decisions rest.

Tourism is of vital importance to the economies of several small island states. Around 65 percent of the small island tourism market is captured by Carribean islands, ensued by Europe, East Asia and the Pacific, Africa and South Asia. The small islands are heavily reliant upon tourism earnings from post-industrial and industrialising countries around the globe whose fascination with island holidays fuels a multimillion dollar industry (Scheyvens and Momsen, 2008).

As a small island developing state (SIDS) of about 1860 square kilometres, Mauritius has been able to craft a robust growth-development track accompanied by a palette of political stability, tough institutional framework and convenient law enactment (Zafar, 2011). Tourists have been flocking over the recent years, spelling a sunny outlook for the future of this industry. This prompted the Mauritius Tourism Promotion Authority to raise the benchmark to at least two (2) million tourists in 2015. The policy, therefore, calls for a robust networking to sell Mauritius as an up-market tourist destination by providing timely information and target new market trends.

In the tourists' decision-making process about a destination lies a crucial component i.e. the length of stay. The growth in total revenues from tourists' related activities is buoyed by longer stays. The purpose of this study therefore, is to identify the factors influencing the length of stay of tourists in Mauritius using state-of-the-art survival model approach in consequence making sustainable recommendations to cement the benefits accruing from the tourism industry.

## 1.2 Objectives of this Study

There are various key motivations for conducting the present research. First,

- Vacation duration remains of paramount significance in tourism management. Hotels' main strategies aim to maximize yield and gross operational profit, reduce fixed costs and maintain high rates of occupancy. The Sun Resorts group includes star-stubbed hotels like Le Touessrok, Sugar Beach, La Pirogue and Ambre. Part of their strategy is to attract tourists who wish to stay longer, including repeaters, since the length of stay is influenced by previous visits, particularly during off-peak seasons. As such, it is vital to ascertain which covariates best explain the length-of-stay decision, since the covariates of the length of stay are specific to the destination.

- Second, survival models have proven to be particularly appropriate to the modeling and analysis of duration events. Although the practicality of survival modeling for predicting duration events has been documented (Menezes *et al*., 2008), their application to tourism remains rather small and this leaves much room for an extensive research.

- Third, it is important for policy purposes to investigate how tourists decide the duration of their stay. Tourists who visit Mauritius only for short periods tend to stay along the coastline and visit only the major tourist attractions (River trek adventure and hiking, swimming with dolphins and whale watching, visit Rhumerie de Chamarel distillery and the seven coloured earth, interaction with lions and cheetahs, Tyrolienne, sky diving). Longer-stay tourists (business, conference, visit friends and relatives, holidays, pilgrimages, shopping, health and medical care, migration possibilities), by comparison, are likely to visit a greater range of attractions, and generate more diverse economic, social and environmental impacts. Families with children also tend to stay longer, in order to save money in the overall package.

- Fourth, the research on length of stay aims to present relevant policy implications. Our main aim is to identify those variables which are relevant for policy implications. Studies have so far focused on the impact of demand-side factors such as the socio-economic characteristics, with little attention given to motivational and psychological characteristics. Ignoring those factors can lead to omitted-variable bias problems. We next consider the supply-side characteristics of the destination as suggested by Martínez-Garcia and Raya (2008). The influence of actual trip experiences (e.g.

natural, social, historical characteristics of the island) is studied. With a rise in the number of arrivals, the tourism sector must also take its share of responsibility and meet the growing demands. Concomitantly, this adheres to the sustainable tourism concept.

- Finally, a representative sample of the visitors' population will be scrutinized. The research takes into consideration the sample selection bias which is a recurrent issue faced by many researchers as it affects both internal and external validity. As per Cuddeback *et al*. (2004), selection bias occurs because non-participation is rarely random and thus, a 100% participation rate is unrealistic. Ignoring it in duration models will bias parameter estimates in an unknown direction, rendering conclusions drawn from them at best tenuous.

## 1.3 Scope for Policy Recommendation

In order to have a comprehensive prospect of tourism demand in Mauritius, this study will explore and uncover the following research questions:

- Does Mauritius achieve its potential in tourism industry?
- What are the challenges awaiting the Mauritian tourism industry?
- What are the parameters that have statistically significant effects on the length of stay?

The results shed light on tourism demand for the length of stay in Mauritius. Using micro-level evidence, we explore the problems faced by the sample of tourists and investigate how the determinants could influence their length of stay. The findings set out policy options and propose relevant mechanisms which can help various stakeholders such as the Government of Mauritius, the Mauritius Tourism Promotion Authority and hotel sector.

## 1.4 Outline of the Report

The report is structured as follows:

Chapter 2 reviews the literature on tourism demand and the various determinants influencing the length of stay in a small island developing state such as Mauritius.

Chapter 3 analyses the tourism sector in Mauritius.

Chapter 4 reviews the econometric estimation techniques used are also discussed in this chapter.

Chapter 5 sets out the sampling strategy, survey methods and methodology adopted in this study.

Chapter 6 provides preliminary findings

Chapter 7 presents the findings on the tourism demand for the length of stay using survival analysis.

Chapter 8 concludes and provides policy recommendations and relevant mechanisms for length of stay which will be helpful for various stakeholders in Mauritius.

# CHAPTER 2: LITERATURE SURVEY

## 2.0 Introduction

This section provides an exploration of previous literature to pin down the actual theoretical gaps. The literature review attempts to provide an overview of length of stay in tourism at different destinations. It started at the beginning of the project and was updated throughout the research period. The influence of socio-economic in addition to various trip attributes, travel information, tourism economic impact, sustainability practices and psychological factors on the length of stay of tourists is reviewed. Their effects are debated below.

## 2.1 Theoretical Review

### 2.1.1 Length of Stay

The concept of length of stay while seemingly an imperative focal point in tourism and economic studies both from a theoretical and practical viewpoint, has garnered massive attention since the number of days' tourists reside at a specific destination is expected to influence the aggregate expenditure. Tourism demand stretches to several paradigms such as tourist arrivals, tourist expenditure, travel exports and length of stay (Menezes et al, 2008). As the main tourism industry's input to growth (Bresson and Logossah, 2008), the length of stay can be defined as the amount of time that the tourist spends at a given destination (De Oliveira Santos *et al*., 2014). It is subject to those tourists who materialise at least one overnight stay at a specific location, precluding the time spent on transport. Conversely, excursionists exclude stay-overs while planning to visit a destination. One of the key elements in a decision-making process, it is one of the riddles unlocked by tourists while planning or taking their trip (Decrop and Snelders, 2004).

The major issue requires to be dealt in vacation decision-making coupled with where to go, how and what to do, with whom and how much to disburse (Decrop and Snelder, 2007). Previous research posits that an extended length of stay in a destination will pull tourists towards a wider collection of activities, thereby stimulating the overall expenditure, sense of affiliation and utility (Menezes et al, 2008). Studies of tourism management reflect that travel duration is a major yardstick for hotels envisaging to maintain high occupancy proportions by luring maximum number of tourists (Alegre and Pou, 2006; Barros, 2008). As an important indicator, it also creates leeway for strategic policy and business implications for tourism destinations (Tussyadiah and Pesonen, 2015).

### 2.1.2 Theory of Consumer Behaviour

The economic justification of consumer behaviour assumes that in the event of a consumption choice decision, the consumer will aspire to optimise his utility subject to budgetary constraints and therefore act rationally based on prices and income (Varian, 1987). The neoclassical theory of consumer behaviour outlines the utility function as the optimal satisfaction derived from the consumption of goods and services However, this conventional economic model refutes the inclusion of social psychological factors. By the same token, economic consumer theory underpins the consumer's demand for a good or service to be the upshot of an amalgamation of consumer preferences and possibilities. Tourism demand is a 'derived' demand and that utility is derived from the characteristics of services rather than from the service itself. In line with the work of Peypoch *et al*. (2015), applying this theory to tourism travel implies that travel characteristics could be employed as inputs in investigating the individual length of stay (which has expenditure attributes), alongside income and price.

Disegna and Osti (2016) postulate that consumers rank goods and services according to particular combinations to reap the highest utility at a certain budget. The separability assumption of the utility function tracks independence among groups of commodities. Concomitantly, tourists start by allocating their budget between activities and other commodities, move on to apportion the budget to a certain destination and eventually among the goods and services offered by the destination. The substantial tourism literature on visitors' expenditure at the macro level portrays economic consumer models, taking into account the fact that heterogeneity and diversity of individual consumer behaviour are warranted. Per se, the survival model under scrutiny for this study is based on the theory of consumer behaviour developed by Lancaster (1966), to treat the demand for tourism as derived. Despite the prevalence of literature in tourism studies and the somewhat restricted research in tourism dealing directly with various determinants, few studies have employed a broader qualitative ensemble of factors affecting the length of stay.

### 2.1.3 Motivation and Socio-economic factors

Among the travel information characteristics, motivation has garnered huge interest in the sphere of travel and tourism (Alen et al, 2010). Push and pull factors have explained the desire to travel and choice of destination respectively, therefore motivating potential tourists to pursue tourism experiences. The broad dimensions of this model propose that people travel because of their own intrinsic forces or needs and are pulled by a certain destination's attributes or perceptions. Propounded by Crompton (1979), tourist motivation has been decomposed into push motives which are internal resting on the need to satisfy disequilibrium or pull motives to arouse the travel yearning. Examples of push factors comprise namely escape, self-exploratory, relaxation, prestige, relapse, kinship-enrichment and social interaction examples of pull factors may take in novelty and education.

Most of the literature concurs on incorporating socio-demographic variables in the model. Many studies chose to include expenditure, price, loyalty and trip motivation (Peypoch et al, 2012; Gokovali et al, 2007; Menezes et al, 2008). Socio-demographic factors include age, gender, nationality, residential area, educational level and income level. Several studies posit that age has a positive relationship with the length of stay (Barros and Machado, 2010; Barros et al, 2010; Martinez-Garcia and Raya, 2008). However, when it attains a certain level, the relationship becomes negative for reasons such as health. A priori the length of stay increases post-retirement age, given the amount of free time available to them (Fleischer and Pizam, 2002). Other reasons include waiving of social obligation (work, family) and carrying discretional income than other groups (Alegre and Pou, 2003).

### 2.1.4 Trip Attributes and Satisfaction

Among the appealing trip attributes cited in the literature, the most prominent ones include security, climate, shopping areas, places of historical/aesthetic interest. By the same token, tourist satisfaction or utility stems from the tourists' experiences in an expected destination. Satisfaction can be defined as "the tourist's emotional state after his holiday trip". In fact, the level of satisfaction pertains to the tourist's sentiment regarding the outcome of a holiday destination and will enhance the probability of the tourist visit the same destination anew, will supply favourable reviews about the destination and therefore construct a positive image (de Menezes et al, 2009). In a tourism setting, the degree of satisfaction has been linked to different components such as hotels, cruises, gastronomy and tour guides without overlooking an umbrella of destination factors such as culture, scenery, entertainment and native environment. Besides, the onset of the information era associated with the exponential growth of the internet has evoked dynamic search to facilitate travel planning. Tourists use the internet, external source of information, to sift through their information needs (Cho et al, 2014). The impact of the internet on tourism cannot be overlooked, given that it is a power house of information (Salman and Hasim, 2012). From the literature, it seems that studies on the external source of information such as the internet and the length of stay are lacking.

### 2.1.5 Crowding-Out Effect

Conceptually, the crowding-out effect traces back to the assumption of Harrod-Domar model which stipulates that the tourism quota is constant over time in a mixed economy. The development of tourism can interfere with the evolution of other sectors under the hypothesis of resource-scarcity. Basically, the term crowding-out effect which has been borrowed from macroeconomics[1] transpires when an economic activity represses another economic activity. The crowding-out of non-tourism businesses (alternative investments in remaining sectors) can become susceptible to the over-arching development of tourism projects subject to limited resources (Candela and Figini, 2012). Like all drivers of economic development, tourism is accompanied by both desirable and undesirable effects. Tourism can influence domestic residents in numerous ways, particularly through the purchase of a service or on-the-spot presence. The influx of tourists will cause crowding (Mathieson and Wall, 1990). Further justification reckons that if the number of tourists exceeds a country destination's tourism capacity in the absence of additional sustenance facilities in time, the excess demand may crowd out the diverse international tourists and stir up a disorder of tourism. Developing larger-scale tourism can trigger crowding out domestic traditional sectors such as fishing, agriculture and mining (Su et al., 2012).

### 2.1.6 Sustainability Incentives

In the advent of rapid globalisation, environmental awareness and sensitivity have turned to be increasingly larger issue in today's marketplace and tourism destinations have not been left unaffected. Small island developing states (SIDS) and those islands in support of small communities represent sensitive cases both for environment and development. Ecologically vulnerable, their minor size, limited resources, geographic dispersion and alienation from markets, place them at an unfavourable position economically and prevent economies of scale. The ocean and coastal environment constitutes a valuable development resource (UNCED, 1992).

[1] In macroeconomic models of aggregate demand, when a certain type of expenditure, for e.g. public expenditure, drives down another demand component, for e.g. private investment, the effectiveness of an expansionary fiscal policy tends to narrow sharply.

As the World Tourism Organisation defines: “sustainable tourism development meets the needs of present tourists and host regions while protecting and enhancing opportunities for the future”. Broadly demarcated, it comprises the fulfilment of economic, social and aesthetic needs while upholding cultural integrity, significant ecological processes, biological diversity and life support systems. The gist of sustainable tourism development is to optimise the positive impacts of tourism, improving and conduct them into the right direction and to curtail the negative impacts (Cernat and Gourdon, 2007). However sustainable tourism has been the subject of several controversies and some researchers resorted to the term sustainable development in tourism. Highly essential to the research on sustainable tourism, the length of stay is indubitably useful in forecasting tourists' duration and in gauging the stress on domestic resources engendered by tourism activity, deemed to be a key issue in the context of carrying capacity analysis. The latter embraces the applied version of sustainable tourism and pertains to the maximum number of individuals who can use an area without modifying its physical features or hampering its quality (Saarinen, 2006; Menezes et al., 2008).

The adoption of sustainable tourism practices and the positive environmental attitude of tourism customers supplies greater avenues for enhancing the environmental performance of hotels. In line with this justification, nature-conscious tourists are more likely to patronise an accommodation which lays emphasis on environmental attitude. Hotels embracing “greening” and environmental programmes in turn are rewarded through international goodwill and longer stays. Conversely, hotels reluctant on adopting environmental-friendly measures may face the peril of declining support or consumer pressure to invest more in eco-friendly products (Masau and Prideaux, 2003). Studies on the environmental practices and initiatives by businesses with respect to tourists have barely been the point of interest in the tourism sustainability literature, increasing the need for further exploration (Mensah et al., 2013).

### 2.1.6 Theories on Personality Traits

In retrospect, a silent revolution has been shaping in personality psychology. The hunt for a scientifically gripping taxonomy of personality factors had begun in the 1970s but its sources can be traced back to the era of Aristotle (Goldberg, 1992). Accordingly, initial explorers in the sphere, Borgatta (1964a, 1969b) revealed similar five structures in their lines of research. Later, the “Big-Five” factors had been categorized into the following: Factor I; Extraversion; Factor II, Agreeableness; Factor III, Conscientiousness; Factor IV, Neuroticism and Factor V, Culture.

However, Factor V has been newly reinterpreted as Intellect (Digman and Takemoto-Chock, 1981; Peabody and Goldberg, 1989) and replaced by Openness-to-experience (McCrae and Costa, 1987). Based on the various findings and protagonists of the big five personality model, the framework is a sympathetic attempt to classify the individual differences as depicted below.

**TABLE 1: THE OCEAN MODEL**

| | |
|---|---|
| **1. Openness** | *(sensitivity, inclination for variety, fantasy)* |
| **2. Conscientiousness** | *(vigilance, efficiency, sense of organisation)* |
| **3. Extraversion** | *(enthusiasm, zeal, assertiveness)* |
| **4. Agreeableness** | *(warmth, optimism, tact)* |
| **5. Neuroticism** | *(anxiety, envy, moodiness, frustration)* |

**SOURCE: AUTHOR'S COMPUTATION**

From Table 1, the five-factor model (FFM) of personality or OCEAN is a meticulous structure of personality traits with respect to the five planes: Openness to experience, Conscientiousness, Extraversion, Agreeableness and Neuroticism. The five-factor model is an account of the trait concept, which has been the overriding example in the European personality psychology, while its assumptions have been embraced by many scholars to seek consistent and persistent individual differences (Thomae, 1989; MrCrae and John, 1992). The OCEAN model has a variety of implications for the tourism behaviour. It can supply a strong and succinct portrayal of an individual's personality (George, 2002).

Firstly, openness-to-experience relates to the various facets of an individual's personality. It conveys intellect and imagination. Typical open-to-experience individuals report high levels of intellectual curiosity. Contrarily, a low-slung level of openness is related to a preference for acquaintance, uncomplicatedness and closure. These individuals portray themselves as conventional, conservative, conformist and behaviourally unyielding (McCrae and Costa, 1987). Secondly, conscientiousness exemplifies an individual's sense of responsibility and, planning and organisation behaviour. Pertaining to other features such as perseverance, diligence and goal-oriented, conscientiousness might be the most significant trait. Low scorers tend to get distracted from task, renege on plans and succumb to weaknesses.

Third, extraversion reflects characteristics such as gregariousness, assertiveness, motivation and sensation-seeking. Displaying a lower form of self-control, they tend to be animated and highly drawn towards outdoor and community or social activities. These sensation-seeking individuals constantly aim for new, first-of-its-kind and complex experiences (Schultz and Schultz, 2008). Forth, agreeableness denotes the exhibition of empathy towards other individuals. Those displaying a high quotient of agreeableness are caring and not reluctant to live in harmony. In general, they tend to be warm and optimistic (Gabbi and Zanotti, 2011).

In the tourism context, owners should offer those products which suit the personality of the customer to optimise both revenue and satisfaction. Obtaining a holistic approach of the subject in consumer behaviour is thorny, despite the basic importance of motivation and personality in adventure travel (Scheider and Vogt, 2012).

## 2.2 Empirical Review

The literature on survival models in tourism studies is incipient but mostly concentrated in developed countries. Paucity of micro data in developing countries persists to be a key source of concern to conduct such analysis. Lately, several researchers adopted the survival analysis instead of ordinary linear regression analysis. We start by delineating the length of stay studies.

Using survival analysis to model the length of stay of tourists of four nationalities in Bodrum, Turkey, Gokovali et al. (2007) employed two model specifications namely the Cox model and the Weibull model. In their aptitude, the covariates were nationality, age, job type, socio-economic characteristics, package vacation type (full-board, half-board, full-package), past visits, quality, level of hospitality, attractiveness, nightlife, accommodation, image, promotion and publicity and recommendation. Their results suggest a robust positive relationship for covariates such as: nationality (Russian), income, experience, independent tour, timing of reservation and familiarity. It is also found that level of education, level of daily expenditure, number of yearly vacation plans, type of holiday and nationality

(British) all mitigate the probability of staying longer. Overall, most of the variables explain the length of stay.

Martinez-Garcia and Raya (2008) investigated the determinants of the length of stay of tourists who are particularly included in low-cost tourism in Spain using the Cox and log-logistics survival models. They advocate that explanatory variables such as the tourist's nationality, age, level of education, type of occupation, type of accommodation selected, season and destination area significantly affect the length of stay. As the hypothesis of proportionality was left unfulfilled, an accelerated failure time model was adopted. In addition, they note that their data was best fitted under the log-logistic model compared to the Cox model.

Menezes et al. (2008) apply a micro-econometric semi-parametric Cox proportional hazard model to scrutinise the determinants of length of stay of tourists in a small island destination of Azores in Portugal with a view to enhance the efficiency of regional tourism policy. They find evidence of socio-demographic profiles, trip attributes, sustainability practices and destination image attributes explaining the length of stay. Concomitantly, trip attributes flagging repeat visitation rates and type of flight were found to be particularly significant.

Artal et al. (2008) employ a count data model by investigating tourists visiting Murcia in the period of 2002-2006. They estimated the Poisson distribution to model the length of stay. Owing to the presence of over-dispersion in the data, the negative binomial regression is then adopted. Empirical evidence demonstrates that nationality, age, travelling with friends, sun-and-sea, price and income largely influence the length of stay. In terms of policy implications, the covariates call for perusal again to optimise the length of stay, given that real expenditures of tourists are on a declining trend in Spain.

Barros and Correia (2008) investigate the determinants of the length of stay of Portuguese tourists specifically holidaying in Latin America using questionnaire data and catering for heterogeneity and sample selection. The covariates under perusal are budget, destination attributes, socio-demographic characteristics, past visits, temporal constraints and the travel frequency. Alternative duration models are estimated for the sake of comparison: the Cox proportional hazard model, the parametric Weibull model, the Logistics model and the Weibull model with heterogeneity. They end up concluding that budget, age, class, friends, Treserve, Ftrip and expectation improve the chances of extending the length of stay. They conclude that the length of stay hinges on various determinants and is subjective to each tourism destination.

Barros et al. (2010) employ a dataset drawn from a survey conducted to golf tourists in Algarve in 2004 to examine the characteristics associated with the length of stay. Among the various duration models, the Cox proportional hazard model, the parametric Weibull model, a Weibull model with heterogeneity and a Weibull model with sample selection are estimated with the former being chosen, based on the log likelihood statistics. They uncover evidence of positive length of stay related to nationality and education of respondents, age, climate, events and hospitality. They also find that the length of stay is negatively associated to beach, showing bleak interest in the principal attribute of sun and sea resorts.

Machado (2010) study the relationship between the destination image and demand duration, peering particularly in the context of Madeira. They employ a sample of international tourists at the culmination of their vacations by having recourse to a questionnaire survey. Between the parametric Weibull model without sample selection and the Weibull model with sample selection, the latter is estimated with superior performance on the grounds of log-likelihood statistic. The empirical evidence shows that the length of stay is positively associated to age, gender, education, German, wine and previous visit. In line with the concept that general economic prosperity stimulates the length of stay, it also indicates other covariates such as British, Dutch and French, expenditure and quality are negatively related to the duration of stay.

To date, sparse studies apart from the developed countries have investigated the length of stay of global tourists. Gearing towards the African continent, Peypoch et al. (2012) contributes to the literature pertaining to the modelling of length of stay of tourists in Madagascar. They adopted a not-so-explored multivariate fractional polynomial survival model to establish the significant characteristics conditioning the length of stay. They use a dataset extracted from a detailed survey in 2007. The parsimonious results reveal that income, age, gender and education have a significant and positive impact on the length of stay, thereby substantiating most the hypotheses. Destination attributes, security and physical appearance also increase the duration.

Oliveira Santos et al. (2014) incorporate a shared heterogeneity into the duration model to estimate the length of stay of tourists. They use an expansive dataset of 309000 visits in Brazilian destinations. They uncover evidence of the optimum intermediate length of stay to corroborate with the theoretical justifications associated to the joint effects of initial expenditure and diminishing marginal utility of the stay. They find that place of origin is a critical determinant of duration with Asians and Oceanians experiencing longer stays whilst Paraguayans account for the shortest stays. Air travel, rented or personal dwellings, sun and sea and previous visits are noted to increase the length of stay. They also find that shared heterogeneity statistically enhances the explanatory capacity of duration models.

Chaiboonsri and Chaitip (2012) investigate the international tourist consumer behaviour during the period 2010-2011 via tourism demand for the length of stay in India. Their research focuses on estimating the count model including both Poisson regression analysis and negative binomial with duration of stay being influenced by social, economic and environment development satisfaction. Their results demonstrate that 4 of the 24 covariates influence the length of stay. However, their estimation did not accommodate specification of the regression model yielding over-dispersion. To wrap up, the estimated model does not describe the Indian tourism demand suitably.

## 2.3 Hypothesis Development

Based on the above theoretical arguments, different strands of hypotheses were skimmed into the following:

**Hypothesis 1 (Socio-economic characteristics)**

The length of stay is a positive function of individual socio-demographic characteristics such as age, gender, level of education, residential area, income and marital status. Their impacts are undetermined.

**Hypothesis 2 (Nationalities)**

The length of a vacation is a positive function of the nationality of the tourist, reflecting to some extent his/her income.

**Hypothesis 3 (Vacation Expenditure)**

The length of a vacation is a negative function of the individual's expenditure. This is a traditional hypothesis in tourism demand models, in which price, income and budget constraints define the frontier of consumption possibilities for travel. Cost of travel is expected to have similar impact. On-trip expenditure incorporates expenses borne when travelling and at the destination.

**Hypothesis 4 (Previous visits)**

The length of stay is positively influenced by previous visits. Kozak (2001) demonstrated that overall satisfaction and the number of previous visits considerably influences the intention to return, especially in some mature destinations.

**Hypothesis 5 (Accompanying party):**

Tourists with an accompanying party are expected to have higher length of stay as it may be less costly to travel in group than individually. As such they have more income to have relatively greater length of stay.

**Hypothesis 6 (Destination attributes):**

The length of stay is a positive function of a destination's attractiveness in attributes such as hotel quality, casino visits, sun and sea, security, gastronomy, hospitability, nature and climate and shopping malls. The Mauritian government has been investing intensively in brand awareness (e.g., sun, sand, sea and shopping fiestas).

**Hypothesis 7 (Satisfaction):**

The length of stay is a positive function of holiday satisfaction.

**Hypothesis 8 (Sustainability incentives):**

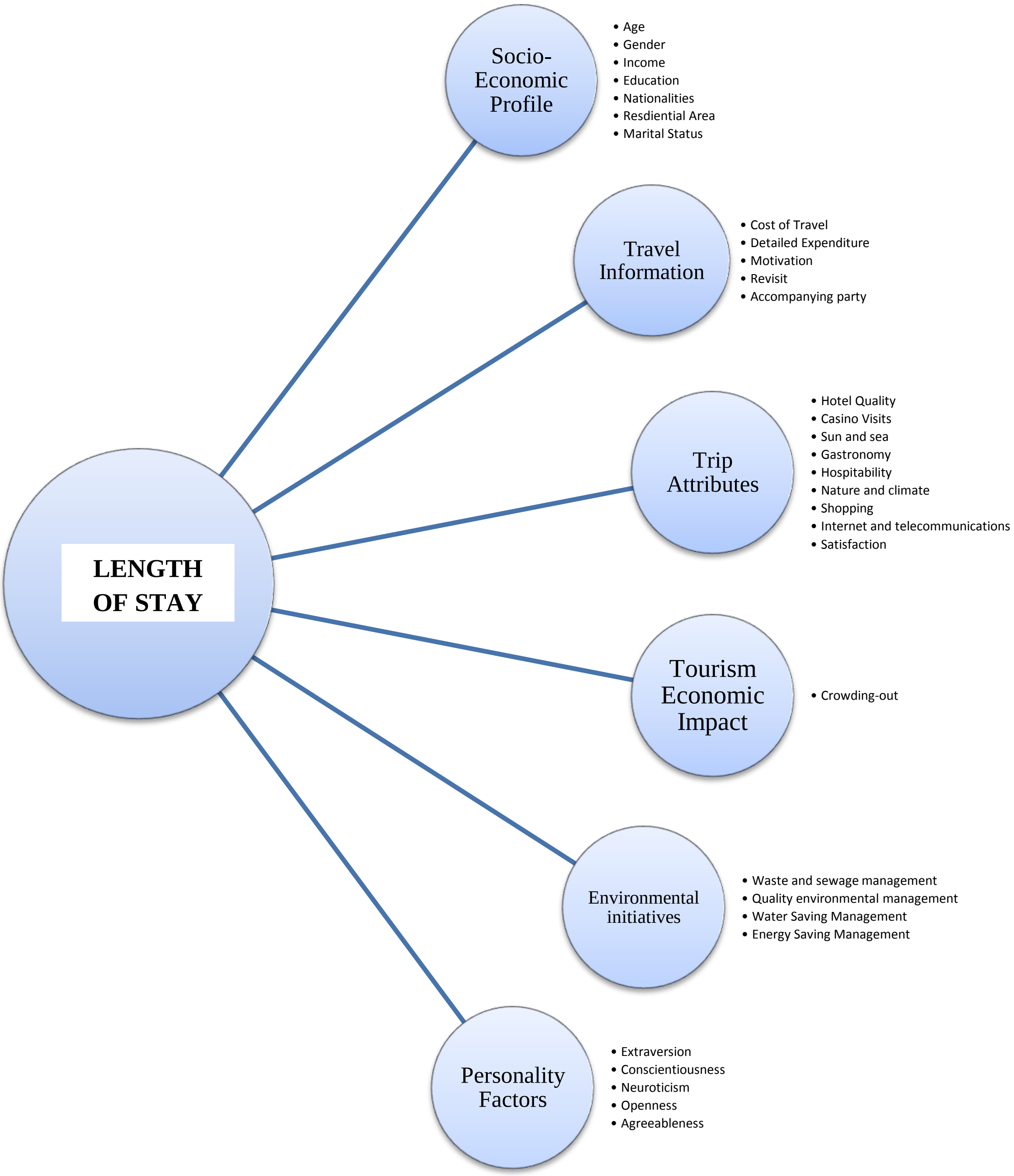
The length of stay can be expected to positively be linked with sustainability practices. The tourist's perception about sustainability practices such as a hotel's waste management, quality of environmental management, water saving management and energy saving management will be pored over.

**Hypothesis 9 (OCEAN Model):**

- The literature has not reached consensus on the impact of openness on length of stay.
- We can expect a negative or non-significant impact of conscientiousness on length of stay.
- A positive effect of extraversion on length of stay is expected.
- A positive effect of agreeableness on length of stay is anticipated.
- A negative impact between neuroticism and length of stay is expected.

What ensues is a collaboration of factors gleaned from the literature in tourism research that is used as the conceptual framework for this study.

**FIGURE 2.1: LENGTH OF STAY FRAMEWORK**




**SOURCE: AUTHOR'S COMPILATION**

## 2.4 Concluding Remarks

Tourism demand studies having recourse to duration models have been done extensively for developed countries and a few developing economies. In the light of the above, the active elements present in tourism should be appraised holistically. The purpose for of this study is to add empirical evidence to the prevailing literature which is rather scarce for small-island-developing countries.

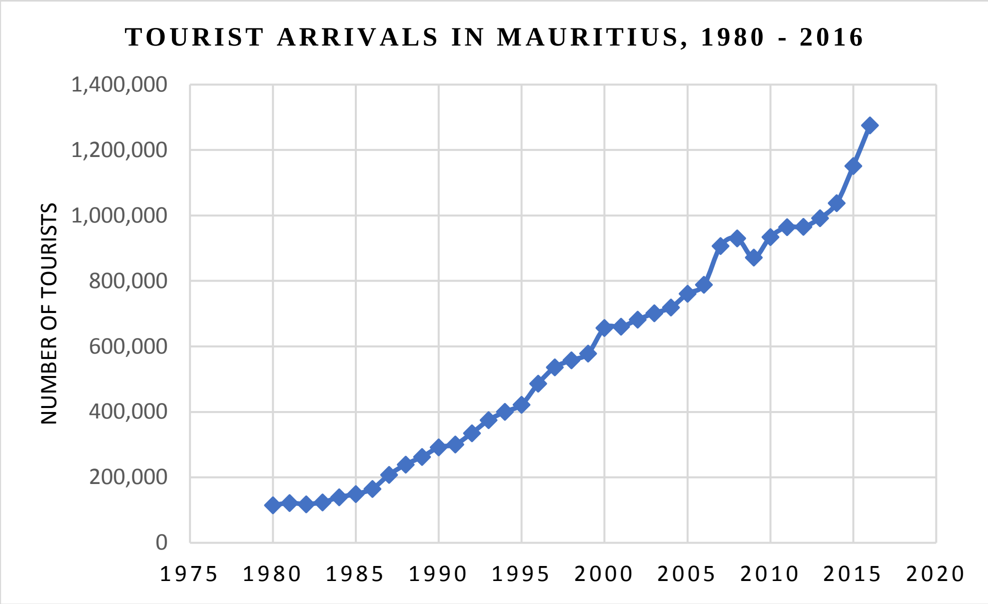
# CHAPTER 3: TOURISM IN MAURITIUS

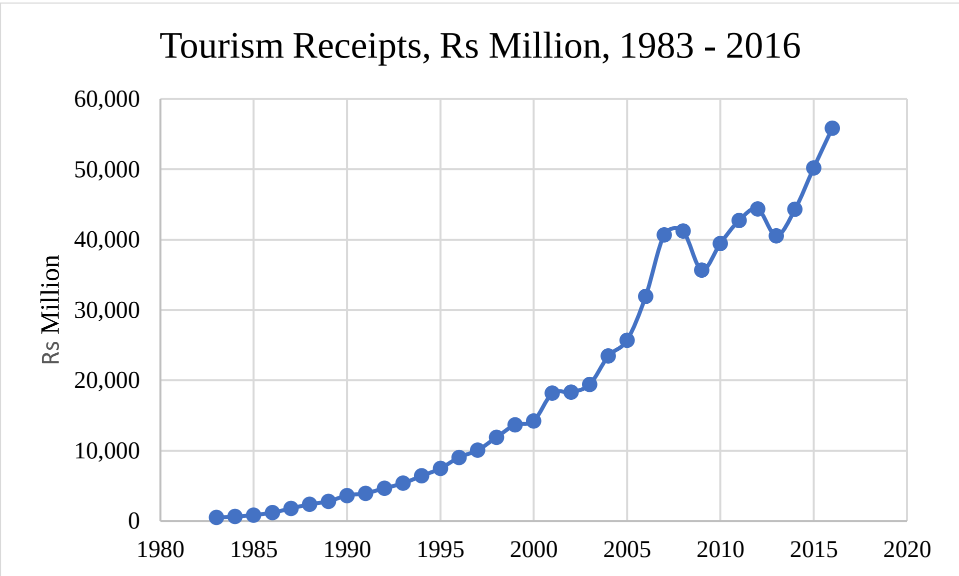
When Mauritius obtained independence in 1968, its economy was based on the sugar sector. By the end of the 1980s and early 1990s, the emergence of new sectors, namely in the field of tourism and financial services surfaced. Mauritius' tourism industry is an important one for the island, and enjoys considerable support from the government. Despite issues on the mainland of Africa, Mauritius is set to enjoy a stable 2015, with an increase in inbound travel of 3.7% from 2014, ending the year at 1.08mn arrivals. The government has supported tourism by launching a number of investment and marketing initiatives in the past, but a major boost has been given by the evolution of the country's financial services industry, which attracts business travel from the continent and countries such as China and India. By becoming a financial hub for the continent and offshore banking in general, there has been a large presence of international banks such as Standard Bank, HSBC, and Barclays, which again helps boost international travel into the country, as well as helping diversify the economy in general. The Mauritian tourism industry can expect to enjoy a stable few years between 2016 and 2019, with average annual growth of 3.3% forecast between 2016 and 2019 leading to a total arrivals figure of over 1.2mn.

Mauritius is well-known by holiday-makers across the globe as an up-market travel destination. Since the early 1980s, tourism has been a key engine of economic growth and development in shaping the domestic landscape. As one of the most important pillars of the Mauritian economy, the tourism sector has recorded robust performances over the precedent years.

Figures 1 to 2 below exhibit the upward trend in aggregate tourist arrivals and tourism earnings while table 1 below shows the distribution of tourist arrivals in Mauritius. Some salient figures about the Mauritian Tourism Sector are shown in the following figures.

FIGURE 1: TOURIST ARRIVALS IN MAURITIUS



Average Length of Stay of tourists, 1998 - 2015

**TABLE 1: DISTRIBUTION OF TOURIST ARRIVAL, BY REGION**

| Country Region | 1998 | 2000 | 2002 | 2004 | 2006 | 2009 | 2010 | 2011 | 2012 | 2013 | 2015 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Africa | 8.2 | 8.0 | 7.9 | 7.4 | 7.7 | 8.1 | 7.7 | 7.8 | 7.2 | 7.5 | 8.1 |
| Europe | 10.6 | 10.2 | 10.4 | 10.7 | 11.1 | 10.9 | 10.9 | 10.2 | 11.2 | 11.0 | 12.1 |
| Asia | 7.6 | 8.6 | 7.7 | 7.5 | 6.8 | 6.5 | 7.3 | 7.3 | 6.8 | 6.2 | 9.1 |
| America | 10.3 | 9.7 | 12.0 | 11.5 | 9.6 | 8.3 | 8.1 | 9.9 | 9.3 | 6.9 | 12.6 |
| Oceania | 11.7 | 10.5 | 13.2 | 9.9 | 10.9 | 11.2 | 11.5 | 10.2 | 9.9 | 7.5 | 13.4 |
| All Countries | 9.9 | 9.6 | 9.6 | 9.7 | 9.8 | 9.7 | 9.6 | 9.3 | 9.5 | 9.2 | 10.6 |

# CHAPTER 4: SURVIVAL ANALYSIS

## 4.1. Introduction to Survival Models

The history of an individual, a government and a machine can be described by a sequence of events. This history includes, for instance, people graduating from university, finding a job, getting married, having children and retiring, a government is being formed, passing new bills, facing steep budget deficits, etc. Such events can refer to the failures and repairs of a machine that runs for production (Van den Berg, 2001, Thrane, 2012). Survival analysis answers questions such as why some patient suffering from a particular type of cancer live longer than other patients, why some tourists tend to stay shorter at a specific destination in comparison to other travellers, why some governments experience lengthy strikes compared to other countries, and why a machine’s life span is longer than another one.

Survival data are generated by “failure time process” of units such as individuals, governments, and machines, which are observed at some specific time and are at risk of transitioning to a new state at any given point in time (Van den Berg, 2001, Lancaster, 1992). Appropriately, survival analysis is a useful tool to understand how a set of explanatory variables can cause variations in the time at which an event may occur. An event in survival models refers to a change or transition from one state to another. For example, an event is time that a state at war transitions to peace, hospitalized person takes to be discharged, and true to this study, the time a tourist takes to leave the destination country and returns home.

In biostatistics, the application of survival analysis deals with estimation and observation of time it takes for lab animals to die due to specific diseases, time for a woman to give birth, and length of life, which all experience transitioning to a new state (Van den Berg, 2001). In engineering, survival analysis is known as "reliability analysis" or "failure time analysis", which models the time it takes for machines or equipment to break down (Murthy et al., 2004). For example, reliability engineers engage duration models to estimate the cost of oil spills by a specific set of explanatory variables such as climate condition, proximity to shore, the availability of facilities.

Social sciences being a different field, employs survival analysis to identify the duration of mortality, fertility, life expectancy, unemployment, business cycles, inflation, war, strike and so forth. Recently empirical analysis of survival models, also known as duration models, has become very popular among econometricians, when time is the variable of interest. In tourism studies, researchers are interested in the time spent in a specific destination and the effect of socio-economic variables on length of stay at the destination (Barros and Machado, 2010, Martínez-Garcia and Raya, 2008, Thrane, 2012, Gokovali et al., 2007)

The core of survival analysis is time. Hence this method is the best suited for modeling and analyzing the duration events such as length of stay (Thrane, 2012, Barros and Machado, 2010). Applying standard techniques such as linear regression to data, which generated under the “failure time process”, can cause severe problems such as bias and the inadequacy in information. The reasons behind such problems are discussed in the following section.

Survival analysis data have some notable specifications which make them incompatible with traditional multiple linear regression techniques such as Ordinary Least Square. First specification is related to censored data in duration models which Ordinary Least Square method is not able to distinguish between censored and uncensored data (Cleves et al., 2010). Survival data analysis is based on following the subject over time until “the change in state” of the subject of interest during the observation time is noticed. If the subject does not experience the change in state, this will be considered right-censored data. Left censoring occurs when one does not observe the start of the event. That means censored data embarks insufficient and incomplete information (Cleves et al., 2010, Lindeboom and van der Klaauw, 2014).

Second, applying regression models to survival data only gives the mean duration (Lindeboom and van der Klaauw, 2014), while one may be interested in the effects of socio-economic variables on probability of leaving the destination country.

Besides, Ordinary Least Square may produce negative predicted values, which has no meaning for the LOS. Duration models do not predict negative values for the dependent variables, which show the duration until the occurrence of event. More precisely, time-to-state-transition is always positive in survival data while the linear regression can predict negative values (Cleves et al., 2010, Thrane, 2012).

**TABLE 2: TERMINOLOGIES IN TOURISM DEMAND SURVIVAL ANALYSIS**

| Terminology in survival analysis | Terminology in tourism demand | Description |
|---|---|---|

| | Concept | |
|---|---|---|
| Event | Event | The event of interest is leaving Mauritius. |
| Survival state | Survival state | The state referring to staying in Mauritius. |
| Failed state | Left state | The state referring to leaving Mauritius. |
| Time-to-failure (survival time) | Time-to-leave (staying time) | The time during which a tourist stays in Mauritius. In this study, survival time is denoted by LOS, i.e., length of stay |
| Survival Probability | Staying Probability | The probability that a tourist stays in Mauritius for a certain time under a given set of explanatory variables. |
| Failure probability | Leaving Probability | The probability that a tourist leaves Mauritius before a certain time under a given set of explanatory variables. |
| Hazard rate (instantaneous risk) | Leaving rate (instantaneous risk of leaving) | The probability that a tourist leaves Mauritius slightly after the time he or she has spent in Mauritius |
| Mean time to failure | Mean time to Leave | The average time that tourists stay in Mauritius, also known as mean survival time. |

It should be noted that all the aforementioned terms are discussed considering the effects of explanatory variables. The following sections discuss the underlying concepts of survival analysis from a mathematical viewpoint.

A survival analysis problem is in threefold. First, the beginning event should be well-defined. Second, a scale or method for measuring time should be available. This could be conventional methods such as minutes, days, months or years. And third, the ending event should also be known.

## 4.2 Survival Time Data: Some Notable Features

There are basically four main types of sampling process providing survival time data:

First, we have the stock sample. In this case, data collection is based upon a random sample of the individuals that are currently in the state of interest, who are typically (but not always) interviewed at some time later, and one also determines when they entered the state (the spell start date).

Second we consider the inflow sample. Data collection is based on a random sample of all persons entering the state of interest, and individuals are followed them until some pre-specified date (which might be common to all individuals), or until the spell ends. For example, when modelling the length of spells unemployment, one might sample all the individuals who began the spell.

The third process is the outflow sample. Data collection is based on a random sample of those leaving the state of interest, and one also determines when the spell began. Per se, the sample would consist of individuals leaving Mauritius.

The final methodology is the population sample. Data collection in this case may be based on a general survey of the population (i.e. where sampling is not related to the process of interest), and respondents are asked about their current and/or previous spells of the type of interest (starting and ending dates).

The longitudinal data may be collected from three main types of survey. First, administrative records, for example information about unemployment spells may be derived from the database used by the government to administer the benefit system. Second, cross-section sample survey, with retrospective questions can be accounted.  Third, panel and cohort surveys, with prospective data collection are another means of data collection. In this case, the longitudinal information is built from repeated interviews on the sample of interest at a number of different points in time. At each interview, respondents can be asked about their current status and changes since the previous interview.

### 4.2.1 Nature of Data: Censoring and Truncation Mechanisms

The above data collection methods provide data which can differ across individuals. The nature of information spells may be very diverse and has key implications for survival study.

#### 4.2.1.1 Censored Data

Censoring occurs when we are not able to observe the event-time. An intrinsic characteristic survival data is the possibility for censoring of observations that is, the actual time until the event is not observed. Because the response is usually a duration, some of the possible events may not yet have occurred when the period for the data collection has terminated. Hence, the event time is not completely observed. Such observation is said to be censored. In the presence of censoring survival data consists of pairs of observations ($\tau_i$, $\delta_i$) where *i* denotes the subject ranging from 1 to *n*. The presence or absence of censoring is denoted by $\delta_i$ where,

$$\delta_i = \begin{cases} 1 = \text{Ending Event Observed} \\ \\ 0 = \text{Observation censored} \end{cases}$$

The survival time is denoted by $\tau_i$, the time till the event if the ending event is observed, and the time censoring occurs if the observation is censored. Thus, the survival data sample contains *n* pairs of survival times and censoring indicators. There are two types of censoring: First, right-censoring occurs when the time-to-event for an individual is not observed. Consider the following an example. Second, left censoring occurs when the beginning event is not observed exactly. Left censoring exists, but is typically not much used in data analysis.

#### 5.2.1.2 Truncated Data

Truncated survival data are those for which there is a systematic exclusion of survival times and where the sample selection effect depends on survival time itself. Truncation can be defined as a period over which the subject was not observed but is, a posteriori, known not to have failed. The statistical difficulty that the truncation causes is that had the respondent failed, he or she would have never been observed. There are three kinds of truncation.

First, consider the figure below which diagrammatically illustrates left truncation.
Left truncation usually arises because we come across a subject who came at risk some time ago. It is the case when only those who have survived more than some minimum amount of time are included in the observation sample. Left truncation is also known as: delayed-entry or stock-sampling with follow-up. If one sample from the stock of persons in the relevant state at some time *t*, and interviews them some time later, then persons with short spells are systematically excluded. The spell is assumed to known in this case, but the subject's arrival is only observed some later date, hence "delayed entry."

### 4.2.2 Survival Distributions

We use the terms survival or duration analysis and duration model in their most general sense. In fact we do not concentrate on duration but rather on the survival and hazard rates.

#### 5.2.2.1 The Survival Function

The analysis of the survival data is based on the use of special modeling distributions of Kalbfleisch and Prentice (2002). Survival analysis is cast in a language all its own. It is a generic term for a group of models which characterizes a probability distribution of a random variable. Let *T* be a non-negative random variable denoting the length of time until an event of interest occurs, i.e. the time-to-failure event. In most areas of econometrics it is customary to describe T by its cumulative distribution function (c.d.f.). The c.d.f. gives the probability that a unit will fail before a certain time t, which is denoted as:

$$F(t) = P\left(T < t\right),\ 0 < t < \infty$$

which specifies the probability that the random variable T is less than some value t. *F(*t) specifies, say, the probability that a spell of unemployment will last no longer than t weeks.

The probability density function (p.d.f) is obtained by differentiating the c.d.f. and is given as:

$$f\left(t\right) = \lim_{\Delta t \to 0} \frac{p(t \le T \le t + \Delta t)}{\Delta t}$$

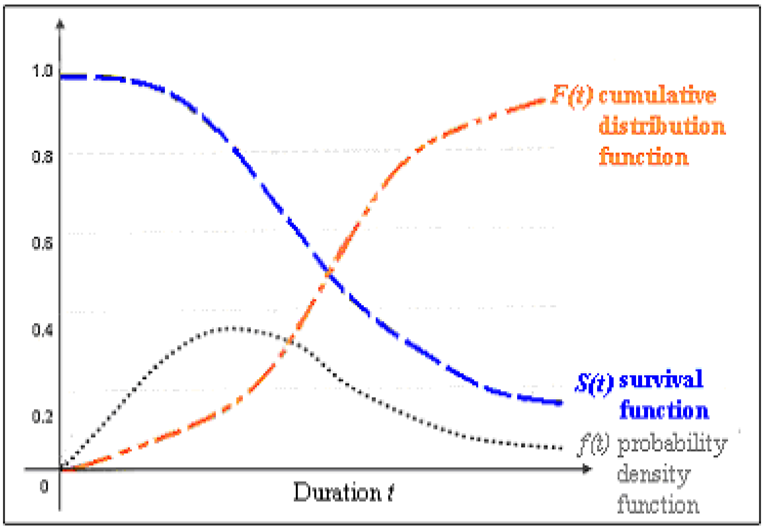
provides an equivalent view. The density function $f(t)$ is a limiting probability that failure will occur between *t* and $\Delta t$. There is no conditioning on the event not happening before time *t*. This is simply the instantaneous unconditional risk of the event happening after time t. However, instead of referring to *T*'s p.d.f., $f(t)$, or c.d.f., $F(t)$ = P($T < t$), survival analysts rather talk about T's survivor function $S(t)$ or its hazard function $h(t)$. The survival function, also called the survivorship or survivor function, is the reverse of the c.d.f. of T:

$$S(t) = P(T < t)$$
$$= 1 - F(t)$$
$$= P(T \geq t)$$

The above equation shows the probability the ending event occurs past time t, i.e. the event time for the random variable T occurs past t. For instance, S(7) is the probability that an individual survives longer than 7 time units, while F(7) is the probability that an individual survives no longer than 7 time units. The survivor function thus reports the probability of surviving beyond time t. As $t \to 0$, $S(t) \to 1$ and $t \to \infty$, $S(t) \to 0$. The survivor function is a monotone, non-increasing function to time. Thus, it is always a decreasing function of time.

The three above-mentioned functions are sketched in Figure 2

FIGURE 2: **MAIN DISTRIBUTIONS**




Source: Computed

### 4.2.3 Hazard Function

The hazard function $h(t)$ is defined as the instantaneous probability the ending event occurs just past time $t$, given that the event had not yet occurred at $t$. It gives the probability or the direct risk that the occurrence of an event can be expected in a interval (usually small) between $t$ and $\Delta t$, provided that the event has not yet occurred until this point in time.

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \le T \le t + \Delta t \mid T \ge t)}{\Delta t}$$

$$h(t) = \frac{f(t)}{S(t)}$$

The hazard function $h(t)$ is also known as the conditional failure rate, the intensity function, the age-specific failure rate, the inverse of the Mills' ratio, and the force of mortality. All three functions above (i.e. c.d.f., p.d.f. and survival function) are associated with $h(t)$.

The hazard rate can vary from zero i.e. no risk at all to infinity i.e. the certainty of failure at the instant. There is a one-to-one relationship between the probability of survival past a certain time and the amount of risk that has been accumulated up to that time. The hazard rate measures the rate at which risk is accumulated. $h(t)$ indicates the "proness to failure" or "risk" of a unit a unit after $t$ has elapsed. It is not a probability as it can take value greater than 1. If we want to work with probabilities, we must use the survivor function.

Due to its definition, the interpretation of the hazard rate requires some care. Hazard rates actually represent latent intensity variables of transition from one state to another, rather than probabilities in a narrow sense (Schneider 1991). The higher the value, the quicker the transition from, say, state A to state B takes place on average. The cumulative hazard function (c.h.f.) is defined as:

$$H(t) = \int_0^t h(t).dt$$

$$\therefore H(t) = \int_0^t \frac{f(t)}{S(t)} = -\int_0^t \frac{1}{S(t)} \left\{ \frac{\partial}{\partial t} S(t) \right\} .\partial t = -\ln\left\{ S(t) \right\}$$

The c.h.f. $H(t)$ measures the accumulated risk or hazard over time 0 and time $t$. *A*s a result,

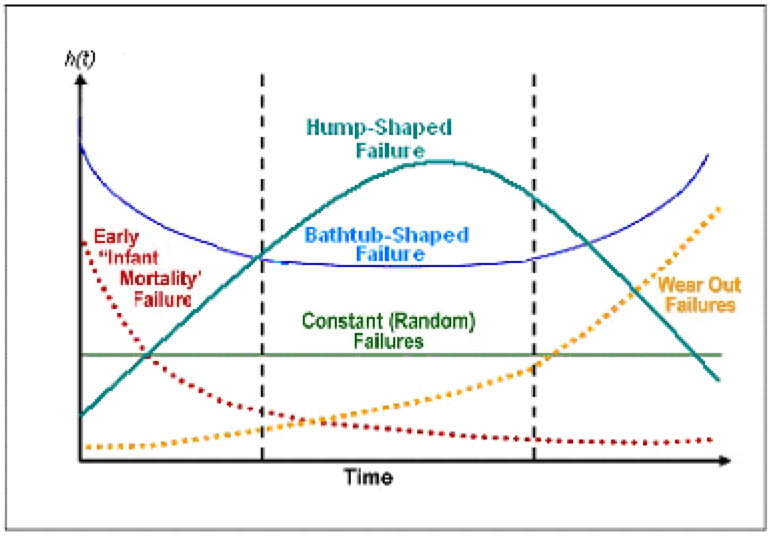
$$S(t) = e^{\{-H(t)\}}$$
$$F(t) = 1 - e^{\{-H(t)\}}$$
$$f(t) = h(t)e^{\{-H(t)\}}$$

$H(t)$ is also called the integrated hazard function. Usually, we require that $\lim_{\partial t \to \infty} H(t) = \infty$ to be able to impose a non-defective distribution of failure times, i.e., $S(\infty) = 0$.

### 5.2.3.1 The Course of the Hazard Function

The shape of the hazard function provides a characterization of the underlying stochastic process. If $\partial h(t) / \partial t > 0$ the process is said to exhibit positive duration dependence, while if $\partial h(t) / \partial t < 0$ the process is said to exhibit negative duration dependence.

**FIGURE 3: COMMON FORMS FOR THE HAZARD FUNCTION**




Source: Computed

The five main courses of $h(t)$ as depicted are:

(i) Constant failure rate: The instantaneous conditional risk of failure is the same, no matter longer the subject has survived up to that point. This kind if hazard behaviour defines the exponential model of survival. Some electronic components may approximate this behaviour. It is also referred to as random-failure rate.

(ii) Increasing failure rate: This kind of behaviour is typical of something that wears out. For e.g. the last part of human life follows this behaviour. As time passes, there's increasing risk of the event.

(iii) Decreasing failure rate: This behaviour occurs whenever there is a wear-in period of adjustment. For e.g., new cars are suggested to have this pattern, the risk of failure decreases as the vehicle is driven, at least earlier in life. The future looks better. This rate is also called early-failures or infant-mortality rate.

(iv) Bathtub-shaped rate: It usually represents the hazard rate of a subject over its lifecycle. For e.g., a car, human life, etc., where they are subject to different rate over time.

(v) Hump-shaped rate: This occurs when the hazard rate is increasing early and eventually starts to decline.

### 4.2.4 Duration Models v/s Linear Regression Models

A model specification of an OLS can be given as: $t_i = \beta_0 + \beta_1 x_i + \ldots \beta_k + \varepsilon_i$, where $\varepsilon_i \sim N(0,\sigma^2)$. Traditional linear models cannot properly account for nonlinearity of longevity data. The distribution of survival data is extremely skewed and often unknown. It is unreasonable, in the case where an event has an instantaneous risk of occurring which is constant over time. Per se, the distribution of time would follow an exponential distribution. Substituting for a more reasonable distribution assumption for $\varepsilon_i$, is the essence of survival analysis. Moreover, right-censored observations and TVC can be incorporated without great conceptual problems. There are two classes of duration models: the proportional hazards (PH) and accelerated failure time (AFT) models. PH models are usually specified as:

$$h(t|\, x_i, \beta) = \lambda_0 exp(\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k)$$
$$h(t|\, x_i, \beta) = \lambda_0(t) exp(\beta_0 + x' \beta_k)$$

, where $h(t|\, x_i, \beta)$ is the hazard for individual i. $\lambda_0(t)$ is the baseline hazard, describing a vector of covariates x′ for an individual and depending on time only. $\beta_k$ is the vector of regression coefficients. An exponential function $exp(\cdot)$ was chosen simply to avoid the problem $h(\cdot)$ ever turning negative. It is termed *proportional hazards* since the hazard subject *i* faces is multiplicatively proportional to the baseline hazard i.e. their ratio is constant over time. The hazard rate of an individual *i* is a fixed proportion (ratio) of for say, individual *j* with different personal attributes. The hazard ratio (HR) is given by:

$$\frac{h_i(t)}{h_j(t)} = \left\{ \exp \beta_1 \left( x_{i1} - x_{j1} \right) + \exp \beta_2 \left( x_{i2} - x_{j2} \right) + \cdots + \exp \beta_k \left( x_{ik} - x_{jk} \right) \right\}$$

AFT[2] models model the log of duration as a function of explanatory variables, while PH models are primarily concerned in estimating the hazard rate. AFT models are given as:

$$\text{Log}_e(T_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_j x_{ij} + \varepsilon_i$$

$T$ is the duration time, $\beta$'s are the coefficient, $x_i$'s are the covariates for $i$ individuals and $\varepsilon_i$ is the error term which needs to be specified.

### 4.2.6 Main Survival Models

Survival analysis makes use of three modelling classes that is the non-parametric, semi-parametric and parametric models.

### 4.2.7 Non-Parametric Models

Non-parametric follows the philosophy of "*letting the dataset speak for itself*" and makes no assumption about the hazard rate. The effects of covariates are not to be modeled. The comparison of the survival experience is done at a qualitative level across values of the covariates. It is labeled "non-parametric" because it is based solely on the data and does not require the estimation of any parameters. Indeed, empirical methods are "……. *less efficient than parametric methods when survival times follow a theoretical distribution and more efficient when no suitable theoretical distribution are known*", Lee (1992). It is suggested to use non-parametric methods before attempting to fit a theoretical distribution.

#### 4.2.7.1 Kaplan-Meier Product Limit Estimators

KM method estimates the survivor function $S(t)$ i.e. the probability of survival past time $t$ (or the probability of failing after $t$). First we shall assume that we have a random sample from the population of spells, and allow for the right censoring (but not truncation). Let $t_1 < t_2 < t_3$ ……

[2] The reason for such appellation is that the effect of the covariate is multiplicative on the time scale. In other words, the effect of the covariate is said to "accelerate" survival time (Hosmer and Lemeshow, 1999). Moreover, a negative sign on the coefficient using the AFT metrics implies a shortening in the duration by some value per unit change in the covariates while for PH metrics, a negative coefficient implies an increase in the hazard rate.

$< t_j < \ldots\ldots < t_k < \infty$ represent the failure times, where $k$ is the number of distinct failure time in the data set. From the data, we can also determine the following quantities:

(i) $d_j$: the number of persons observed to "fail" (make a transition out of the state) at $t_j$.

(ii) $m_j$: the number of persons whose observed duration is censored in the interval $[t_j, t_{j+1}]$, i.e. still in state at time $t$ but not in state by $t$+1.

(iii) $n_j$: the number of persons at risk of making a transition immediately prior to $t_j$ , which is made up of those who have a censored or completed spell of length $t_j$ or longer:

$$n_j = (m_j + d_j) + (m_{j+1} + d_{j+1}) + ... + (m_k + d_k)$$

Table 4.1: Example of data structure

| Failure Time | Number of Failures | Number of Censored | Number at Risk of Failure |
|---|---|---|---|
| $t_1$ | $d_1$ | $m_1$ | $n_1$ |
| $t_2$ | $d_2$ | $m_2$ | $n_2$ |
| $t_3$ | $d_3$ | $m_3$ | $n_3$ |
| - | - | - | - |
| - | - | - | - |
| - | - | - | - |
| $t_j$ | $d_j$ | $m_j$ | $n_j$ |
| - | - | - | - |
| - | - | - | - |
| $t_k$ | $d_k$ | $m_k$ | $n_k$ |

Source: Jenkins (2004)

The proportion of those entering a state who survive to the first observed survival time $t_1$, $\hat{S}$ ($t$), is one minus the proportion who made a transition out of the state by that time, where the latter is estimated by the number of exists divided by the number who at risk of transition:

$$\frac{d_1}{d_1 + m_1} = \frac{d_1}{n_1} \qquad \text{----- (3.3.7.1b)}$$

The proportion of surviving to the second observed survival time $t_2$ is $\hat{S}$ ($t_1$) multiplied by one minus the proportion who made a transition out of the state between $t_1$ and $t_2$.

$$\hat{S}(t) = \prod_{j|t_j \le t} (1 - \frac{d_j}{n_j})$$

The above equation is the empirical survival time $t_j$, where $n_j$ is the number of individuals at risk at time $t_j$ and $d_j$ is the number of failures at time $t_j$. The product is over all observed failure times less than or equal to $t$. The KM estimate is given by the product of one minus the number of exits divided by the number of persons at risk of exit i.e. the product of one minus the "exit rate" at each of the survival times. $\hat{S}(t)$ is also the generalized maximum likelihood estimates of $S(t)$ over all possible distributions. We can derive an estimate of the failure function $\hat{F}(t_j) = 1 - \hat{S}(t_j)$, and the integrated hazard function $\hat{H}(t_j)$, since,

$S(t) = \exp[-\int_0^t \theta(u).\partial u] = \exp[-H(t)]$

$\Rightarrow \hat{H}(t_j) = -\log \hat{S}(t_j)$ ----- (3.3.7.1d)

The standard error (SE) for the KM estimate is given by Greenwood's (1926) formula:

$$\hat{V}ar\{\hat{S}(t)\} = \hat{S}^2(t) \sum_{j|t_j \le t} \frac{d_j}{n_j(n_j - d_j)}$$

These SE are not used for confidence intervals. But, the asymptotic variance of $\ln\{-\ln \hat{S}(t)\}$,

$$\hat{\sigma}^2(t) = \frac{\sum \frac{d_j}{n_j(n_j - d_j)}}{\left\{\sum \ln(\frac{n_j - d_j}{d_j})\right\}^2}$$

, is used, where the sums are calculated over $j|t_j \le t$ (Kalbfleisch and Prentice, 2002). The confidence intervals are calculated as

$$\hat{S}(t)\mathrm{e}^{\left\{\pm z_{\alpha/2}\hat{\sigma}(t)\right\}}$$

, where $z_{\alpha/2}$ is the $(1 - \alpha/2)$ quantile of the normal distribution.

#### 4.2.7.2 Nelson-Aalen Estimators

There exists an alternative non-parametric method for estimating $H(t)$ that has better small-sample properties. We could use the theoretical relation between $H(t)$ and $S(t)$, i.e. $H(t) = -\ln \{S(t)\}$. The estimator is due to Nelson (1972) and Aalen (1978),

$$\hat{H}(t) = \sum_{j|t_j \le t} \frac{d_j}{n_j} \qquad \text{-----}$$

(3.3.7.2a)

, where $n_j$ is the number at risk at time $t_j$, $d_j$ is the number of failures at time $t_j$, and the sum is overall distinct failure times less than or equal to $t$. The standard errors are based on the variance calculation of Aalen (1978),

$$\hat{V}ar\left\{\hat{H}(t)\right\} = \sum_{j|t_j \le t} \frac{d_j}{n_j^2}$$

, and the confidence intervals reported are

$$\hat{H}(t)\mathrm{e}^{\left\{\pm z_{\alpha/2}\hat{\phi}(t)\right\}}$$

, where

$$\hat{\phi}^2(t) = \frac{\widehat{Var}\left\{\hat{H}(t)\right\}}{\left\{\hat{H}(t)\right\}^2}$$

estimates the asymptotic variance of $\ln \hat{H}(t)$ and $z_{1-\alpha/2}$ is the $(1 - \alpha/2)$ quantile of the normal distribution.

### 4.2.8 Non-Parametric Equality Tests of Survivor Functions

The survival equality tests work by comparing at each failure time, the expected versus the observed number of failures for each group and then combining these comparisons over all observed failure times, not at a specific point in time.

#### 4.2.8.1 The "Cox" Test

This term "Cox" test is from Statacorp (1999). It is performed by estimating a Cox regression on *K*-1 indicator variables, one for each of the groups with one of the indicators omitted. The $\chi^2$ test reported is the likelihood-ratio test or the Wald test. It is a variation on the log-rank test.

In the literature, these tests have been popularly used. There are no favoured tests but these can be used as a preliminary test before employing duration regressions.

### 4.2.9 Semi-Parametric Models

The two renowned semi-parametric models are namely the Cox PH and the Piecewise Constant Exponential (PWCE) models. They are referred to as semi-parametric models mainly because the shape of the baseline hazard function is left unspecified.

#### 4.2.9.1 The Cox PH Model

The Cox PH model was invented by Cox (1972) and is one of the most cited paper in survival literature by far. Cox's method does not assume a particular distribution for the survival times, but rather assumes the effects of the different variables on survival are constant over time and are additive in a particular scale.

$$h(t|\, x_i, \beta) = \lambda_0(t) exp(\beta_0 + x' \beta_k) \qquad \text{----- (3.3.9.1a)}$$

The hazard for any individuals $i$ at time $t$ is the product of two components: a baseline hazard function $\lambda_0(t)$ which is left unparameterized and a linear function set of $x'$ fixed covariates. No assumptions about the shape of the baseline hazard $\lambda_0(t)$ over time are made i.e. it could be constant, increasing or decreasing. What is assumed is that, whatever the shape, it is the same for any individuals. The conditional probability that the $i^{th}$ graduate leaves unemployment at time $t$, (given those who could have left at time $t$), is define by the HR. The HR is also the hazard of individual $i$, divided by the sum of all hazards for those individuals $j$ who have not left. Therefore,

$$HR(t|\, x_i,\, \beta) \;=\; \frac{h(t|\mathrm{x_i},\, \beta)}{h(\mathrm{t}|\, \mathrm{x_j},\, \beta)} = \frac{\lambda_0 \exp\left(\beta_0 + x_i'\beta_k\right)}{\sum_{\mathrm{j=i}}^{\mathrm{n}} \lambda_0 \exp\left(\beta_0 + x_i'\beta_k\right)}$$

$$HR(t|\, x_i,\, \beta) \;= \frac{\exp\left(\beta_0 + x_i'\beta_k\right)}{\sum_{\mathrm{j=i}}^{\mathrm{n}} \exp\left(\beta_0 + x_i'\beta_k\right)}$$

The baseline hazard drops out from the calculation. A more technical and detailed treatment is given by Kalbfleisch and Prentice (2002). The major advantage of the Cox specification is that there is no need to make assumptions about the baseline hazard. Any wrong assumption can lead to misleading estimates of $\beta$. However, where the Cox model gains in flexibility, it loses in predictive power. It is very difficult to estimate quantities of interest, such as the median survival time and predicted failure time from the model. Even worse the Cox model simply ignores data when no failures occur in period $t$.

#### 4.2.10 Parametric Models

There are two classes of parametric models: the PH and AFT models. There are basically three types of PH and five types of AFT models (Statacorp, 1999). Note that the notations which are used below follow mainly that of Cleves *et al* (2004).

##### 4.2.10.1 The Exponential Distribution

(i) The Exponential Regression in the PH metric

The baseline hazard being constant, $\lambda_0(t) = 1$. Hence, the model can be specified as:

$h(t|\, x_i, \beta) = \lambda_0(t) exp(\beta_0 + x'\beta_k)$

$h(t|\, x_i, \beta) = exp(\beta_0 + x'\beta_k)$

(ii) The Exponential Regression in the AFT metric

$\ln(T_i) = \beta_0 + x'\beta_k + \ln(\tau_i)$ $\quad$, $\tau_i \sim \exp\{\exp(\beta_0)\}$

The random variable $\tau_i$ is distributed exponentially with mean $\exp(\beta_0)$. The exponential has a purely random failure pattern, or memoryless property. The hazard rate is constant. If we assume a probability of $p$ of surviving for time t, then $\lambda$ is determined by $\lambda = -\log(p)/t$. A small lambda indicates relatively low risk and long survival and vice-versa.

#### 4.2.10.2 The Weibull Distribution

(i) The Weibull Regression in the PH metric

The baseline hazard takes the form $\lambda_0(t) = pt^{p-1}\exp(\beta_0)$. $p$ is the ancillary shape parameter to be determined from the data and $\exp(\beta_0)$ is the scale parameter.

$h(t|\, x_i, \beta) = pt^{p-1}\, exp(\beta_0 + x'\beta_k)$

(ii) The Weibull Regression in the AFT metric

$\ln(T_i) = \beta_0 + x'\beta_k + \ln(\tau_i)$ $\quad$, $\tau_i \sim \text{Weibull}(\beta_0, p)$

, where the c.d.f. is $F(\tau) = 1 - \exp\left[-\{\exp(\beta_0)\tau\}^p\right]$.

The random variable $\tau_i$ is distributed with parameters ($\beta_0$, $p$). When the hazard rate rises monotonically with time ($p > 1$), there is a positive duration dependence in the data. If it falls monotonically with time ($0 < p < 1$), then we have a negative duration dependence. A constant hazard rate ($p = 1$) surmises no duration dependence (i.e. similar to an exponential model).

#### 4.2.10.3 The Gompertz Distribution

The Gompertz PH model assumes a baseline hazard $\lambda_0(t) = \exp(\gamma t)\exp(\beta_0)$. $\gamma$ is the ancillary shape parameter to be estimated from the data and $\exp(\beta_0)$ is the scale parameter.

$h(t|\, x_i, \beta) = \exp(\gamma t)\, exp(\beta_0 + x'\beta_k)$

If $\gamma < 0$, the hazard function is decreasing, while $\gamma > 0$ indicates an increasing hazard rate over time. If $\gamma = 0$, the hazard function reduces to an exponential model. Just like the Weibull, the model is suitable for modelling data with monotonic hazard rates.

#### 4.2.10.4 The Log-Normal Distribution

The log-normal AFT models assume the following:

$\ln(T_i) = \beta_0 + x' \beta_k + \ln(\tau_i)$ , $\tau_i \sim$ log-normal($\beta_0$, $\sigma$)

, where the c.d.f. is $F(\tau) = \Phi\left[\dfrac{\ln\tau - \beta_0}{\sigma}\right]$.

The random variable $\tau_i$ is distributed with parameters ($\beta_0$, $\sigma$) i.e., it follows a standard normal distribution with mean 0 and standard deviation $\sigma$. The attractive feature of the log-normal model is the non-monotonic hazard function i.e. it increases initially and then decreases. Thus, the hazard function is inversed U-shaped.

#### 4.2.10.5 The Log-Logistic Distribution

The log-logistic AFT model is given as:

$\ln(T_i) = \beta_0 + x' \beta_k + \ln(\tau_i)$ , $\tau_i \sim$ log-logistic($\beta_0$, $\gamma$)

, where the c.d.f. is $F(\tau) = 1 - \left[1 + \left\{\exp(-\beta_0)\tau\right\}^{\frac{1}{\gamma}}\right]^{-1}$.

The random variable $\tau_i$ is distributed with parameters ($\beta_0$, $\gamma$) i.e., it follows a standard normal distribution with mean 0 and standard deviation $\pi\gamma/\sqrt{3}$. The log-logistic distribution bears a close resemblance to the log-normal one. If $\gamma < 1$, the log-logistic hazard rises and then falls. If $\gamma \geq 1$, then the hazard is monotonically decreasing.

#### 4.2.10.6 The Generalized Gamma Distribution

The generalized gamma model is given as:

$\ln(T_i) = \beta_0 + x'\beta_k + \ln(\tau_i)$ , $\tau_i \sim \text{gamma}(\beta_0, \kappa, \sigma)$

The c.d.f. is given as $F(\tau) = \begin{cases} I(\gamma, u), if\, \kappa > 0 \\ \Phi(z), if\, \kappa = 0 \\ 1 - I(\gamma, u), if\, \kappa < 0 \end{cases}$

, where $\gamma = |\kappa|^{-2}, z_0 = sign(k)\{\ln(\tau) - \beta_0\}/\sigma, u = \gamma \exp(\sqrt{\gamma} z_0)$, $\Phi(\bullet)$ is the standard normal cumulative distribution function, and $I(a, x)$ is the incomplete gamma function. The random variable $\tau_i$ is distributed with three parameters ($\beta_0$, $\gamma$, $\lambda$). The hazard function is highly flexible in shape, such as including the possibility of an inversed U-shaped or a "bath-tub" shaped hazard. If $\gamma = 1$, we have the Weibull model. If $\gamma = 1$, $\lambda = 1$, we have the exponential model. With $\gamma = 0$, the log-normal model results. And if $\gamma = \lambda$, then one has the standard gamma distribution. The generalized gamma can thus be quite useful for testing model specification.

### 4.2.8 Frailty Models

In multivariate models so far taken into consideration, differences among individuals were assumed to be capture using observed independent variables (e.g. the *x* vector). It is extremely rare to come across data that measured all relevant variables, both *t* and *x*, without error. Some variables are omitted in these models because they may be simply intrinsically unobserved (e.g. ability, motivation, etc.). Such mixed survival models are usually referred to as "frailty models". The term "frailty" was first suggested by Vaupel et al (1979) in the context of mortality studies and Lancaster (1979) in the context of duration of unemployment.

If these unobserved effects are ignored, then Jenkins (2004) suggests three outcomes. First, the "no-frailty" model will over-estimate the degree of negative duration in the hazard (i.e. under-estimate the degree of positive duration dependence). Second, the proportionate response of the hazard rate to a change in a regressor *k*, will no longer constant (i.e. $\beta_k$ in the models without

unobserved heterogeneity), but declines with time. Third, one will get an under-estimate of the true proportionate response of the hazard to a change in a regressor *k* from the no-frailty-model $\beta_k$. A frailty model can be classified into two types, namely the over-dispersion or heterogeneity (unshared frailty) and random-effects (shared frailty) models.
Both semi-parametric and parametric (PH and AFT) models can be extended to include these unobserved individual effects.

**4.2.11.1 Unshared Frailty Models**

An unshared frailty PH model (as per Cleves *et al*, 2004) usually defines the hazard function to be:

$$h(t|\, x_i,\, v_i) = v_i\, h(t|\, x_i)$$
$$h(t|\, x_i,\, \beta) = \lambda_0(t)\; v_i\, exp(\beta_0 + x'\,\beta_k)$$

, where $\nu_i$ captures the unobserved specific effects. The effect, $\nu_i$ is known as a frailty and serves to represent that individuals in the population are heterogeneous due to factors that remain unobserved. In other words, it also describes the unobserved or random effects which multiplicatively influence the hazard of each individual or a group of individuals. The random variable $v_i$ is assumed to have the following properties:

(i) $\nu_i > 0$.
(ii) $E(v_i) = 1$ (unit mean, a normalization required for identification).
(iii) Finite variance $\sigma^2 > 0$.
(iv) Distributed independently of *t* and *x*.

The frailty component explains a portion of the variance which cannot be accounted for by the model containing only fixed effects. A simple transformation $s_i = \log v_i$ permits inclusion of the frailty term in the term $exp(\beta_0 + x'\,\beta_k)$ of the regression model. If *z* represents the incidence vector for the random effect $s = \{s_i\}$ the mixed survival model can be re-written as:

$$h(t|\, x_i,\, \beta,\, z) = \lambda_0(t) exp(\beta_0 + x'\,\beta_k + z's)$$

Two renowned distributions are assumed for the frailty term $v$, namely the inverse-Gaussian distribution and the gamma distribution.

The inverse-Gaussian distribution is given below:

$$g(\nu_i)=\left(\frac{1}{2\pi\theta\nu_i^3}\right)^{1/2}\exp\left\{-\frac{1}{2\theta}\left(\nu_i-2+\frac{1}{\nu_i}\right)\right\} ,v_i \sim \text{inverse-Gaussian}(1,\theta)$$

The frailty term $v$ follows an inverse-Gaussian distribution with mean one and variance $\theta$.

The gamma distribution is given as follows:

$$g(\nu_i)=\frac{\nu_i^{1/\theta-1}\exp(-\nu_i/\theta)}{\Gamma(1/\theta)\theta^{1/\theta}} \text{ , } v_i \sim \text{gamma}(1,\theta)$$

The frailty term $v$ follows a gamma distribution with mean one and variance $\theta$.

#### 4.2.11.2 Shared Frailty Models

Shared frailty models are analog to random-effects models. The frailties are no longer observation specific, but are shared across groups of observations. Consequently, this causes those observations within the group to be correlated. The generalization of the PH frailty model can be modified to accommodate shared frailties across groups of observations.

$$h\left(t_{ij}\left|x_{ij},\nu_{ij}\right.\right)=\nu_{ij}h\left(t_{ij}\left|x_{ij}\right.\right)$$

, for data consisting of $n$ groups with the $i^{\text{th}}$ group of $n_i$ observations. The index $i$ denotes the group ($i = 1,\ldots,\ldots,n$), while $j$ denotes the observation within group, $j = 1,\ldots,\ldots,n_i$. The frailties $v_i$ are shared within each group and are assumed to follow either a gamma or inverse-Gaussian distribution. The frailty variance $\theta$ is estimated from the data and measures the variability of the frailty across groups.

# CHAPTER 5: SAMPLING STRATEGY AND SURVEY METHOD

## 5.1 Sampling Strategy

This is a cross sectional study which aims at determining the length of stay of tourists in Mauritius. More specifically, it focuses on tourism involvement in the Mauritian landscape and relates to their overall characteristics, trip attributes, sustainability perspectives and personality traits.

The study is restricted to tourists aged 18 and above and who spent some days in Mauritius. Statistics Mauritius estimates that the target population consists of 1,275,227 individuals.

Information pertaining to the distribution of the target population was made available prior to conducting the fieldwork. Consequently, we used the quota sampling to identify and targeted to interview a sample of 2000 individuals from the target population. It would be difficult to collect data about all the tourists due to time and budget constraint, thus justifying the choice of a sample survey. In a tourism survey setting, one could virtually establish a sampling frame for simple, stratified or systematic random sampling at the departure gate as soon as the last passenger turns up and the boarding procedure has not yet initiated. However, practically it is not feasible as no time would be left for recruitment and data collection. Ideally, airport surveys are usually cross-sectional and non-probability samples (convenience, quota, snowball) are suitable (Bauer, 2015).

The quota sample is obtained by dividing the given population into mutually exclusive segments and to determine the proportions the sub-groups are made of. The same proportion would be used during the sampling process. Individual observations were chosen from the subgroups so that the final sample is a representative of the whole population.

| Continents | Amount in Population | Percentage (%) | Amount in Sample |
|---|---|---|---|
| **Europe** | 734,506 | 57.60 | 1,152 |
| **Africa** | 291,890 | 22.89 | 458 |
| **Asia** | 208,233 | 16.33 | 327 |

| | | | |
|---|---|---|---|
| **Oceania** | 20,071 | 1.57 | 31 |
| **America** | 19,766 | 1.55 | 31 |
| **Others** | 761 | 0.06 | 1 |
| | 1,275,227 | | 2,000 |

**SOURCE: COMPUTAION BASED ON STATISTICS MAURITIUS DATA, 2016**

The central aim of sampling was to ensure the sample is representative of the broader population from which it is drawn.

Quota sampling is a non-probability sampling technique wherein the assembled sample has the same proportions of individuals as the entire population with respect to known characteristics, traits or focused phenomenon. Quota sampling is a viable method because of prevalent diverse characteristics in the population. The members in each group formed have comparable attributes and characteristics. The strata are known as "quota controls" and selected based on their pertinence to the topic of interest (Yang and Banamah, 2013). With quota sampling, the researcher ensures that a quota or specified number of participants is selected with pre-set features. Similar to stratified sampling, the researcher selects subjects according to his suitability and judgement. If the participants in quota sampling are assigned equal probability of inclusion in the sample, quota sampling will be at par with stratified sampling which is a probability sampling method. However, as the selection procedure in quota sampling is left to the interviewers the samples drawn upon the two methods concur only in quota-controls and may diverge in other characteristics (Yang and Banamah, 2013; Singleton and Straits, 1999)

Among the advantages of quota sampling, the latter tend to save on time and easier to carry out as it discards the requirement of having a sampling frame. The quota sample enhances the representation of particular groups within the population, making certain that the strata are not victims of over-representation. Comparisons can be made fairly with the use of a quota sample. However, the drawbacks stemming from this method demonstrate that the sampling error and its origin become difficult to assess. The quota sampling has been heavily criticised by academics as statistical inferences between sample and population are not likely which may trigger the issue of generalisation. Failing to split data into subgroups can eventually lead to sampling overestimation or underestimation. Overall, it can be useful if a good strata-division method is carefully employed. This technique is the best alternative as compared to other non-probability sampling techniques (Na et al., 2017).

Considering the above discussion, the quotas are established for each country of origin, in proportion to the total international tourists' arrivals in 2016 in Table 3. The Statistics Mauritius revealed that the top 10 tourists generated markets in 2016 are France, Reunion Island, United Kingdom, South Africa, Germany, India, China, Switzerland, Italy and Australia. The top and grossing 10 markets are accountable for 80% of the aggregate international tourists' arrival in Mauritius.

*Table 3: Quota Sampling according to country of origin, based on tourist arrivals 2016*

| Country of Origin | Tourist Population | Sample Size |
|---|---|---|
| | **(Tourist Arrivals 2016)** | **(n=2000)** |
| **France** | 271,963 | 427 |
| **Reunion Island** | 146,203 | 229 |
| **United Kingdom** | 141,904 | 223 |
| **South Africa, Rep. of** | 104,834 | 164 |
| **Germany** | 103,761 | 163 |
| **India** | 82,670 | 130 |
| **Peoples' Rep. of China** | 79,374 | 124 |
| **Switzerland** | 36,272 | 57 |
| **Italy** | 31,337 | 49 |
| **Australia** | 18,559 | 29 |
| **Austria** | 16,643 | 26 |
| **Belgium** | 15,675 | 25 |
| **Spain** | 15,304 | 24 |
| **Sweden** | 14,551 | 23 |
| **Madagascar** | 11,740 | 18 |
| **Poland** | 10,126 | 16 |
| **Netherlands** | 10,080 | 16 |
| **United Arab Emirates** | 9,614 | 15 |
| **Russian Federation** | 9,295 | 15 |

| | | |
|---|---|---|
| **USA** | 8,524 | 13 |
| **Czech Republic** | 8,503 | 13 |
| **Denmark** | 6,817 | 11 |
| **Seychelles** | 6,393 | 10 |
| **Canada** | 6,060 | 10 |
| **Korea** | 6,025 | 9 |
| **Malaysia** | 4,628 | 7 |
| **Norway** | 4,589 | 7 |
| **Finland** | 4,415 | 7 |
| **Portugal** | 4,222 | 7 |
| **Turkey** | 3,486 | 5 |
| **Ireland** | 3,441 | 5 |
| **Ukraine** | 3,230 | 5 |
| **Kenya** | 3,185 | 5 |
| **Saudi Arabia** | 3,164 | 5 |
| **Slovakia** | 2,934 | 5 |
| **Brazil** | 2,912 | 5 |
| **Singapore** | 2,840 | 4 |
| **Romania** | 2,813 | 4 |
| **Japan** | 2,655 | 4 |
| **Philippines** | 2,488 | 4 |
| **Hungary** | 2,432 | 4 |
| **Indonesia** | 2,375 | 4 |
| **Other American** | 2,270 | 4 |
| **Zimbabwe** | 2,047 | 3 |
| **Namibia** | 1,715 | 3 |
| **Bulgaria** | 1,707 | 3 |
| **Pakistan** | 1,610 | 3 |

| | | |
|---|---|---|
| **Other European** | 1,561 | 2 |
| **Luxembourg** | 1,534 | 2 |
| **Taiwan** | 1,457 | 2 |
| **Nigeria** | 1,372 | 2 |
| **Hong Kong SAR [4]** | 1,342 | 2 |
| **Mayotte** | 1,189 | 2 |
| **Israel** | 1,161 | 2 |
| **Mozambique** | 1,144 | 2 |
| **Greece** | 1,122 | 2 |
| **Botswana** | 1,088 | 2 |
| **Zambia** | 999 | 2 |
| **Other African** | 979 | 2 |
| **New Zealand** | 939 | 1 |
| **Slovenia** | 916 | 1 |
| **Ghana** | 889 | 1 |
| **Bangladesh** | 844 | 1 |
| **Iran** | 837 | 1 |
| **Comoros** | 800 | 1 |
| **Lithuania** | 792 | 1 |
| **Afghanistan** | 787 | 1 |
| Others & not stated | **761** | 1 |
| **Vietnam** | 729 | 1 |
| **Congo** | 724 | 1 |
| **Morocco** | 674 | 1 |
| **Tanzania** | 625 | 1 |
| **Egypt** | 607 | 1 |
| **Latvia** | 589 | 1 |
| **Nepal** | 587 | 1 |

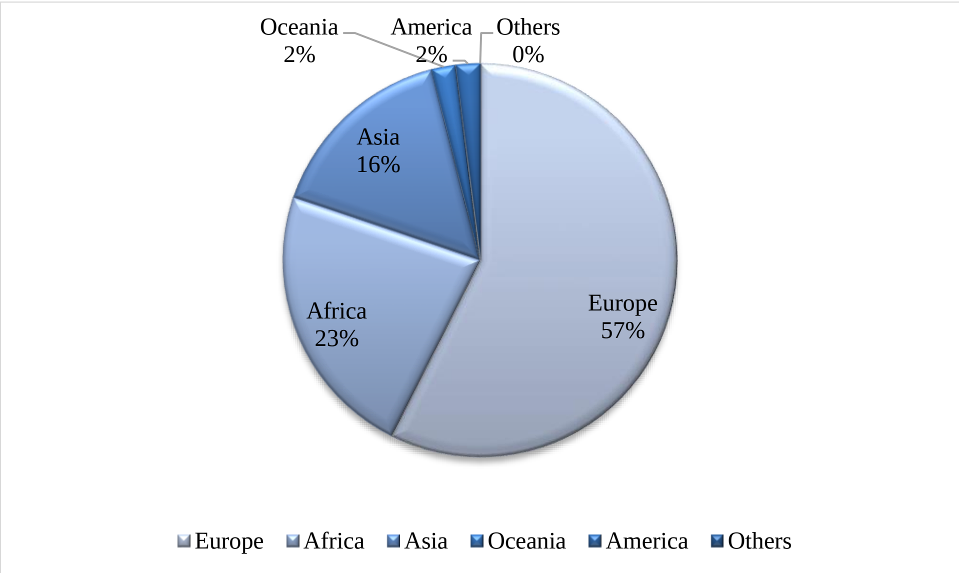
| | | |
|---|---|---|
| **Other Oceanian** | 573 | 1 |
| **Angola** | 505 | 1 |
| **Estonia** | 489 | 1 |
| **Uganda** | 469 | 1 |
| **Thailand** | 466 | 1 |
| **Kuwait** | 461 | 1 |
| **Serbia** | 455 | 1 |
| **Croatia** | 449 | 1 |
| **Belarus** | 422 | 1 |
| **Malawi** | 402 | 1 |
| **Qatar** | 385 | 1 |
| **Tunisia** | 382 | 1 |
| **Lebanon** | 380 | 1 |
| **Sri Lanka** | 379 | 1 |
| **Swaziland** | 370 | 1 |
| **Kazakhstan** | 355 | 1 |
| **Ivory coast** | 350 | 1 |
| **Other CIS** | 322 | 1 |
| **Cameroon** | 321 | 1 |
| **Algeria** | 310 | 0 |
| **Other Asian** | 296 | 0 |
| **Ethopia** | 272 | 0 |
| **Oman** | 245 | 0 |
| **Lesotho** | 223 | 0 |
| **Gabon** | 215 | 0 |
| **Senegal** | 201 | 0 |
| **Rwanda** | 183 | 0 |
| **Bahrain** | 165 | 0 |

| | | |
|---|---|---|
| **Niger** | 164 | 0 |
| **Jordan** | 154 | 0 |
| **Burundi** | 93 | 0 |
| **Sudan** | 79 | 0 |
| **Togo** | 78 | 0 |
| **Benin** | 66 | 0 |
| **Other Middle East** | 61 | 0 |
| **Maldives** | 54 | 0 |
| | **1,275,227** | **2,000** |

Source: Author's Computation

Each quota pertains to every country of origin share as a percentage of respondents. The statistics used have been sought from the Statistics Mauritius database highlighting tourist arrivals for the year 2016. In addition, we can have an overview of the sampling proportion by the main continents including Europe, Africa, Asia, Oceania, America and others. It is noteworthy that France with 21% of total tourist arrivals has been the major source country for Mauritius, followed by the sister island Reunion (12%), the United Kingdom (11%), South Africa (8%) and Germany (8%).

**FIGURE 4 SAMPLING PROPORTION BY CONTINENT**



Source: Author's Computation

From the pie chart above, it becomes crystal clear how tourism in Mauritius is dwarfed and dominated by European countries, followed by African and Asian sources.

### 5.2 Questionnaire Design and Administration

Data was collected by way of drop-off survey at the departure lounge of the Sir Seewoosagur Ramgoolam International Airport. In leisure and tourism studies, the method of questionnaire-based surveys is the widely-used technique to extract information about individual's accounts of their behaviour (Veal, 2006; Brunt, 1997).

To capture accurate information about the length of stay of tourists visiting Mauritius, it was proposed that the data was to be collected at departure points in Mauritius. The survey initiated on 17$^{th}$ January 2017 and concluded on 31$^{st}$ May 2017. It covers 1870 tourists who travelled to Mauritius from different corners of the globe. The sample was designed in such a way as to allow for age, country of origin and educational background of the respondents. The field work was carried out at the departure lounge of the Sir Seewoosagur Ramgoolam International Airport. We were able to discuss with a number of international tourists, but the main instrument was questionnaire, designed as follows:

| Section | Questions and Rationale | Measurement Scale |
|---|---|---|
| **Introduction** | Briefly explains the research topic and guarantees confidentiality and anonymity | |
| **Section A “Personal Profile”** | Gathers data about the socio-economic characteristics, nationalities, expenditure incurred during the stay, number of previous visits and accompanying party of the respondent. | Closed-ended questions, dichotomous and open-ended questions |
| **Section B “Trip Attributes** | The general statements express the drive behind the different trip attributes including the hotel (accommodation) quality, casino visits, sun-and-sea, security, gastronomy, hospitability, nature and climate, shopping, the importance of internet and telecommunications. The last question addresses the level of satisfaction expressed by the tourists in Mauritius. | Five-point Likert Scale |
| **Section C “Sustainability Practices”** | Respondents were mandated to evaluate how far various sustainability practices are important for them to have a high-quality holiday. The environmental initiatives pertain to waste and sewage management, quality environmental management, water saving management and energy saving management. | Five-point Likert Scale |
| **Section D “Personality Traits”** | Questions derived from the OCEAN model | Five-point Likert-Scale |

Source: Author's Computation

Since our target population accommodates those tourists leaving Mauritius by air at the Sir Seewoosagur Ramgoolam International airport as they would be in a better position to express their views based on their experiences with multiple aspects of their stay, the following key points are taken into consideration:

- The country of origin of the tourists
- The main languages spoken.
- Collections of data were extended over afternoons and evenings during the months of February, March and April 2017
- The different assortment of tourism
- Tapping different days of the week, including weekends.

Notably, the interviews are conducted on the 'first available' basis implying that the inbound tourists were approached as they entered the departure lounge after check-in formalities. A short explanation on the purpose of this study was given to each tourist. To facilitate better understanding, the designed questionnaires are available in English, French and Mandarin(Chinese) so that certain tourists are not left out of the targeted population. The language issue is recognised by the United Nations World Tourism Organisation (2005) stipulating the following: "for border survey aimed at non-residents not only is the translation of the questionnaire particularly important but also the language in which the interview is conducted".

The questionnaire was administered using the drop and collect method which allows for an important level of security and reliability during the data collection phase. At the end of each day, the surveys were synchronized with the survey database, thereby updating the survey findings and generating up to date objectives to the interviewers. The average length of the questionnaire is 15 minutes. The questionnaire was translated and administered mostly in English and French, and in some very rare cases the questionnaire was administered in Mandarin Chinese. An experienced and skilful network of interviewers who have been trained to deploy various survey techniques was employed for this assignment. Interviewers were selected based on the following criteria:

- Command of the English and French language

- Prior experience working on similar projects

- Good sense of contact
- Professional presentation/attire
- Good diction
- Availability during the entire duration of the survey from 7 a.m. to 11 p.m. including Saturdays, Sundays and Public holidays.

Those few who declined to cooperate just could not do otherwise because of their late arrival for check-in formalities.

**5.3 Quality Assurance**

The group of interviewers was accompanied by one field supervisor who exerted close control on the fieldwork. This was to allow for remote control and direct verification of the questionnaire during the administration phase while the project managers also closely monitored fieldwork.

**5.4 Respondents' Incentives**

In exchange for taking part in the survey, each respondent was rewarded by a souvenir in the form a metal keyring embossing the map of Mauritius.

# CHAPTER 6: Survey Analysis

## 6.1 An Anatomy of the Response Rate

The questionnaire targeted the tourists who completed their vacation throughout the whole period of observation. Hence, censored data is precluded in the analysis. In total 1870 tourists were interviewed. Respondents with missing information and incomplete fields for study variables were discarded. The final returned and useable questionnaires amounted to 1647, representing about 82.35% of the selected sample which is rather good for such kind of survey. According to Walonick (1993), the response rate is the single most important indicator of how much confidence can be placed in the results of a mail survey. Kanuk and Berenson (1975) argue the higher the response rate, the more accurate the survey. Babbie (1973) advocates a 50% response rate as an appropriate threshold. However, he adds that these are rough estimates guides, they are no statistical basis, and a demonstrated lack of response bias is far more important than response rate.

## 6.2 Descriptive Analysis and Graphical Illustrations

A perusal of the remaining variables of the available dataset is of paramount importance. The above table unveils some of the descriptive statistics. Scale variables, or those that could reasonably be treated as such are summarized using the mean as measure of central tendency and standard deviation (SD) as measure dispersion, whereas categorical variables are reported using counts and percentages. They are crucial, albeit elementary, summary statistics. The average age in the sample is 41 years. Evidently, the sample is coarsely divided into about 39% male and 61% female.

**TABLE 4 SUMMARY STATISTICS FOR CATEGORICAL VARIABLES**
**SOURCE: COMPUTED**

| Variable | Options | Count | Percentage |
|---|---|---|---|
| **Gender** | Male | 1003 | 60.9% |
| | Female | 644 | 39.1% |
| **Educational Attainment** | Primary | 13 | 0.8% |
| | Secondary | 214 | 13.0% |
| | Bachelor | 698 | 42.4% |
| | Masters | 539 | 32.7% |
| | Others | 183 | 11.1% |
| **Country Region** | Europe | 961 | 58.3% |
| | Africa | 380 | 23.1% |
| | Asia | 260 | 15.8% |
| | Oceania | 22 | 1.3% |
| | America | 24 | 1.5% |
| **Residential Location** | Urban | 1010 | 61.3% |
| | Semi-Urban | 427 | 25.9% |
| | Rural | 210 | 12.8% |
| **Income** | Level 1 | 134 | 8.1% |
| | Level 2 | 388 | 23.6% |
| | Level 3 | 477 | 29.0% |
| | Level 4 | 320 | 19.4% |
| | Level 5 | 328 | 19.9% |
| **Marital Status** | Single | 358 | 21.7% |
| | Married | 1062 | 64.5% |
| | Divorced | 73 | 4.4% |
| | Cohabitation | 106 | 6.4% |
| | Others | 48 | 2.9% |
| **Purpose of Visit** | Holiday | 1436 | 87.2% |
| | Business | 74 | 4.5% |
| | Visit friends and relatives | 102 | 6.2% |
| | Conference | 13 | 0.8% |
| | Others | 22 | 1.3% |
| **First-time visit in Mauritius?** | Yes | 639 | 38.8% |
| | No | 1008 | 61.2% |
| **Accompanying Party?** | Yes | 365 | 22.2% |
| | No | 1282 | 77.8% |

On average, each tourist stays in Mauritius for 8 days and spends approximately 3653 USD on various items such as accommodation, food and drinks, transport, shopping, entertainment among others. About 87% of tourists visit the island for holiday purposes while only 4% come on business trips and 6% visit their relatives and friends. The remaining 0.7% and 1% comprise those on conference visits and other purposes respectively. Other purposes included holiday destination weddings, medical tourism and pilgrimage in Mauritius.

**TABLE 5 SUMMARY STATISTICS FOR SCALE VARIABLES**
**SOURCE: COMPUTED**

| Variable | | Mean | SD | Min | Max |
|---|---|---|---|---|---|
| **Length of Stay (days)** | | 8.4 | 4.6 | 1 | 60 |
| **Age (years)** | | 41.2 | 14.6 | 18 | 79 |
| **Travel Expenditure ($)** | | 3653.1 | 2559.0 | 0 | 39000 |
| **Number of people accompanying** | | 1.8 | 7.6 | 1 | 149 |
| **Number of previous visits** | | 2.6 | 5.6 | 1 | 52 |
| **Trip Attributes** | Hotel Quality | 4.3 | 0.9 | 1 | 5 |
| | Casino Visits | 1.4 | 0.9 | 1 | 5 |
| | Sun and Sea | 4.4 | 0.9 | 1 | 5 |
| | Security | 4.1 | 0.9 | 1 | 5 |
| | Gastronomy | 3.2 | 1.1 | 1 | 5 |
| | Hospitability | 4.4 | 0.9 | 1 | 5 |
| | Nature and Climate | 4.3 | 1.0 | 1 | 5 |
| | Shopping | 2.3 | 1.2 | 1 | 5 |
| | Internet and Telecommunication | 3.4 | 1.3 | 1 | 5 |
| | Satisfaction | 4.4 | 0.8 | 1 | 5 |
| **Tourism Economic Impact** | Crowding-Out | 3.0 | 0.9 | 1 | 5 |
| **Environmental Initiatives** | Waste and Sewage Management | 4.2 | 1.0 | 1 | 5 |
| | Quality Environmental Management | 4.1 | 0.8 | 1 | 5 |
| | Water Saving Management | 4.1 | 0.9 | 1 | 5 |
| | Energy Saving Management | 4.2 | 0.9 | 1 | 5 |
| **Personality Traits** | Extraversion | 3.6 | 0.6 | 1 | 5 |
| | Conscientiousness | 4.3 | 0.6 | 1 | 5 |
| | Neuroticism | 2.1 | 0.9 | 1 | 5 |
| | Openness | 3.1 | 0.6 | 1 | 5 |
| | Agreeableness | 4.0 | 0.7 | 1 | 5 |

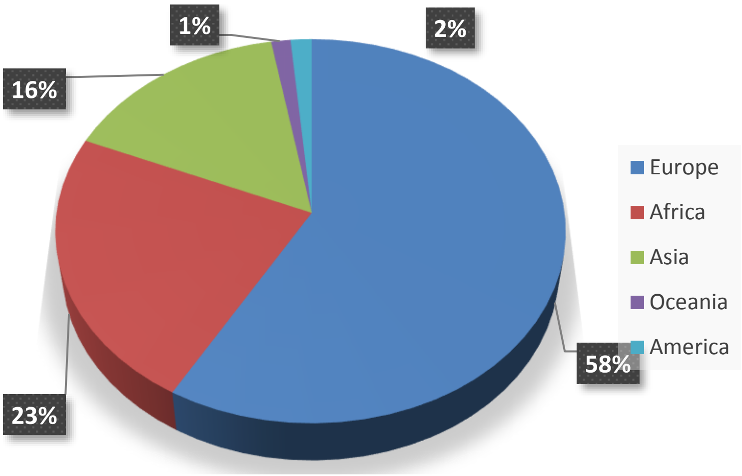
The final sample includes tourist data originating from 51 countries across the globe as shown in the table below. It can be noted that the largest segment came from France (23.9%), followed by United Kingdom (12.3%), Reunion Island (10.8%), South Africa (9.1%), India (7.5%), Germany (7.3%), China (6.6%) and Italy (2.9%).

TABLE 6: ***RESPONSE, BY COUNTRY OF ORIGIN***

| Country of Origin | Frequency | Percent |
|---|---|---|
| Australia | 22 | 1.34 |
| Austria | 8 | 0.49 |
| Belgium | 24 | 1.46 |
| Brazil | 8 | 0.49 |
| Canada | 6 | 0.36 |
| China | 108 | 6.56 |
| Congo | 3 | 0.18 |
| Czech Republic | 12 | 0.73 |
| Denmark | 5 | 0.3 |
| Finland | 8 | 0.49 |
| France | 394 | 23.92 |
| Germany | 120 | 7.29 |
| Ghana | 1 | 0.06 |
| Greece | 3 | 0.18 |
| Hungary | 5 | 0.3 |
| India | 124 | 7.53 |
| Ireland | 5 | 0.3 |
| Israel | 2 | 0.12 |
| Italy | 47 | 2.85 |
| Kenya | 5 | 0.3 |
| Korea | 10 | 0.61 |
| Kyrgyz Republic | 1 | 0.06 |
| Luxembourg | 2 | 0.12 |
| Madagascar | 18 | 1.09 |
| Malawi | 1 | 0.06 |
| Malaysia | 7 | 0.43 |
| Namibia | 3 | 0.18 |
| Netherlands | 17 | 1.03 |
| Norway | 3 | 0.18 |
| Pakistan | 4 | 0.24 |
| Poland | 10 | 0.61 |
| Portugal | 3 | 0.18 |
| Reunion | 178 | 10.81 |
| Russia | 13 | 0.79 |
| Rwanda | 1 | 0.06 |
| Seychelles | 13 | 0.79 |
| Slovenia | 2 | 0.12 |
| South Africa, Rep of. | 149 | 9.05 |
| Spain | 15 | 0.91 |
| Sri Lanka | 2 | 0.12 |
| Sweden | 29 | 1.76 |
| Switzerland | 28 | 1.7 |
| Togo | 1 | 0.06 |
| Turkey | 4 | 0.24 |
| UAE | 2 | 0.12 |
| Ukraine | 2 | 0.12 |
| United Kingdom | 202 | 12.26 |
| United States | 10 | 0.61 |
| Zambia | 2 | 0.12 |
| Zimbabwe | 5 | 0.3 |
| Total | 1,647 | 100 |

**SOURCE: COMPUTED**

**FIGURE 5 DISTRIBUTION OF SURVEY RESPONDENTS BY COUNTRY REGION, 2017**

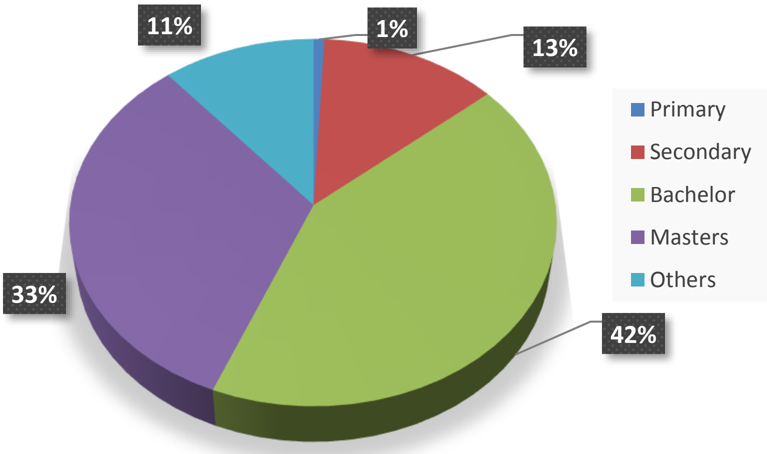



**SOURCE: COMPUTED**

As can be visually seen, the discrepancies between the collected and projected sample are few.

In Figure below, out of all tourists interviewed, 42% declared that the majority completed their Bachelor degree while 33% bagged a Master's degree. It can be further postulated from the sample that 11% and 13% have completed their Primary and Secondary Education respectively whilst the remaining 1% ticked Others. Under the educational attainment, "Others" includes primarily diploma-holders, PhD holders, professional course holders and military staff college holders.

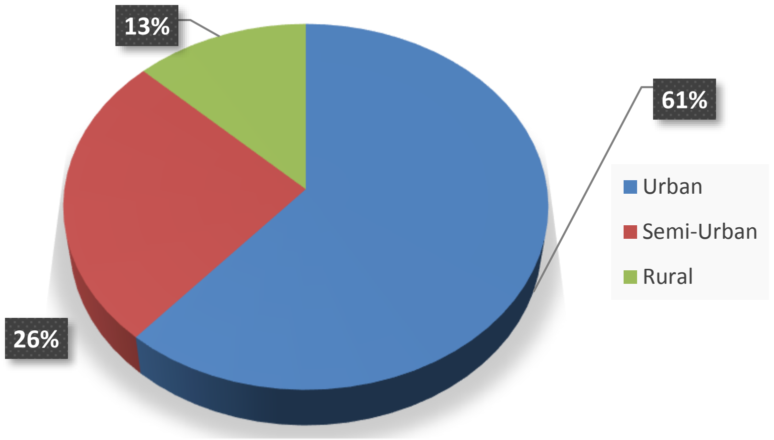
**FIGURE 6 DISTRIBUTION OF RESPONDENTS BY EDUCATIONAL ATTAINMENT, 2017**




**SOURCE: COMPUTED**

Figure 3 below depicts the native region where the tourists emanate. It can be observed that 61% come from urban areas, 26% stem from semi-urban areas whilst the rest belong to rural areas.

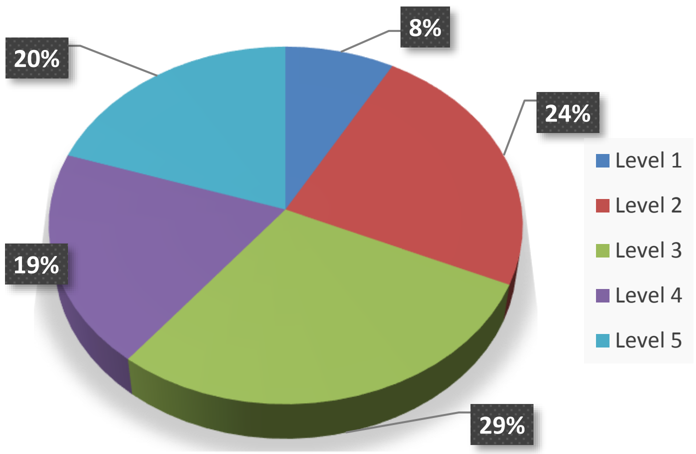
FIGURE 7 DISTRIBUTION OF RESPONDENTS BY REGIONAL LOCATION, 2017




SOURCE: COMPUTED

From the figure below, the bulk of respondents obtain a monthly income of level 3 (between \$2000 and \$3000) while 19% of tourists gain an income of level 2 (between \$1000 and \$2000).

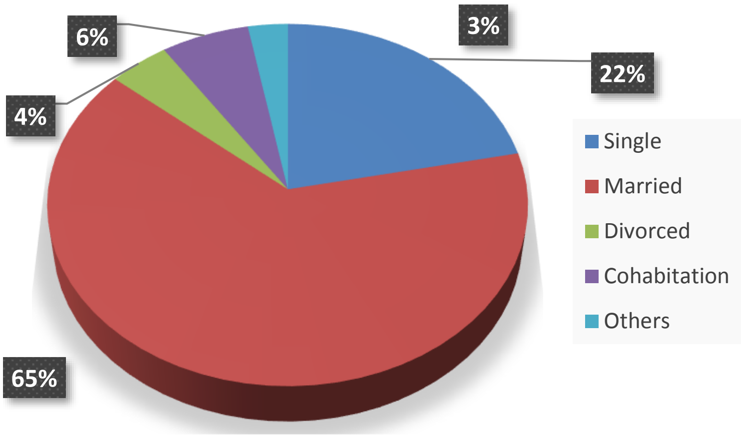
FIGURE 8 DISTRIBUTION OF RESPONDENTS BY INCOME LEVEL, 2017




SOURCE: COMPUTED

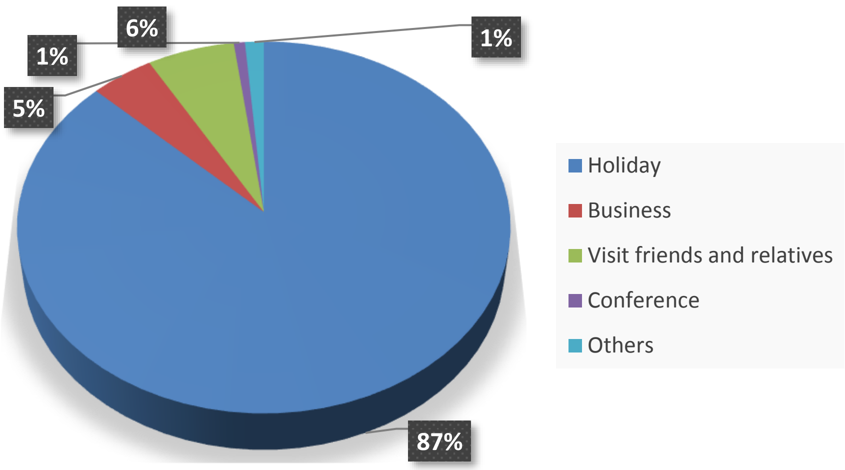
In Figure below, out of all tourists interviewed, 65% revealed that they were married while for 22% showed their status as single. It can be further postulated from the sample that 6% and 4% were in cohabitation phase and divorced respectively whilst the remaining 3% ticked Others. Under the marital status, “Others” included those who got widowed.

FIGURE 9 DISTRIBUTION OF RESPONDENTS BY MARITAL STATUS, 2017



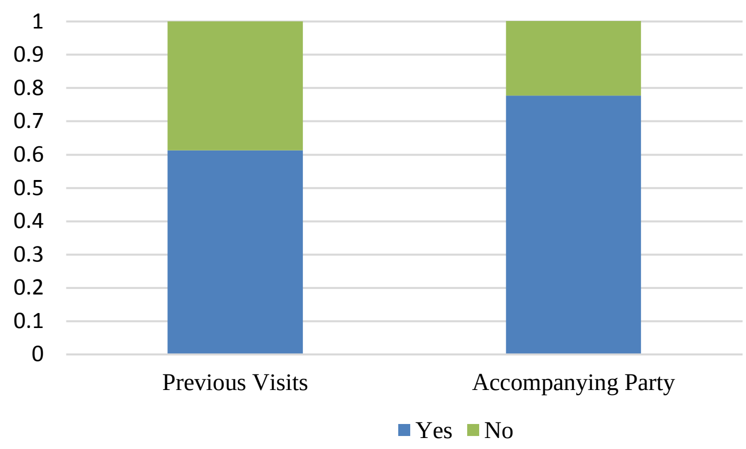

Among the interviewed tourists, almost 87% chose to visit Mauritius for holiday purposes while 6% declared that visiting friends and relatives was the prime drive. Only 5% travelled to Mauritius to conduct business and trade. A minority of 1% rode to Mauritius for conference purposes and the remaining 1% includes other reasons such as medical tourism, pilgrimage during the Maha Shivratri festival and big fat wedding destination.

FIGURE 10 DISTRIBUTION OF RESPONDENTS BY PURPOSE OF VISIT, 2017




**SOURCE: COMPUTED**

FIGURE 11: FURTHER CHARACTERISTICS OF VISITORS




## 6.3 Trip Attributes, Sustainability Practices and Personality Traits

The respondents were asked to rate the following trip attributes to show their approval from 1 to 5 with "1" as "strongly disagree" and "5" as "strongly agree". The following results were yielded.

FIGURE 12: TRIP ATTRIBUTES

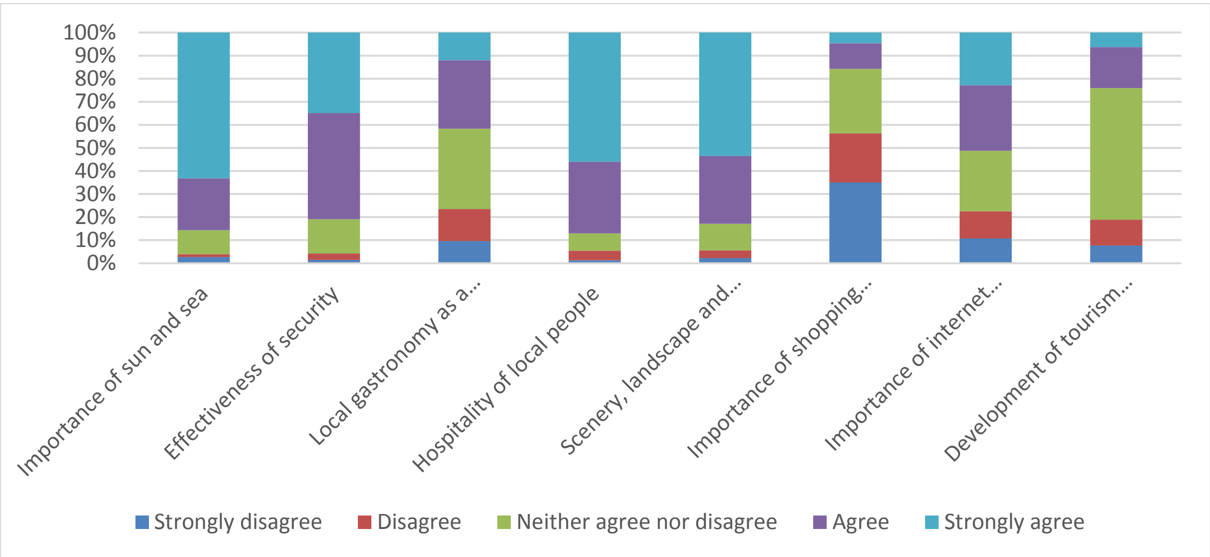


**FIGURE 13: SUSTAINABILITY PRACTICES**

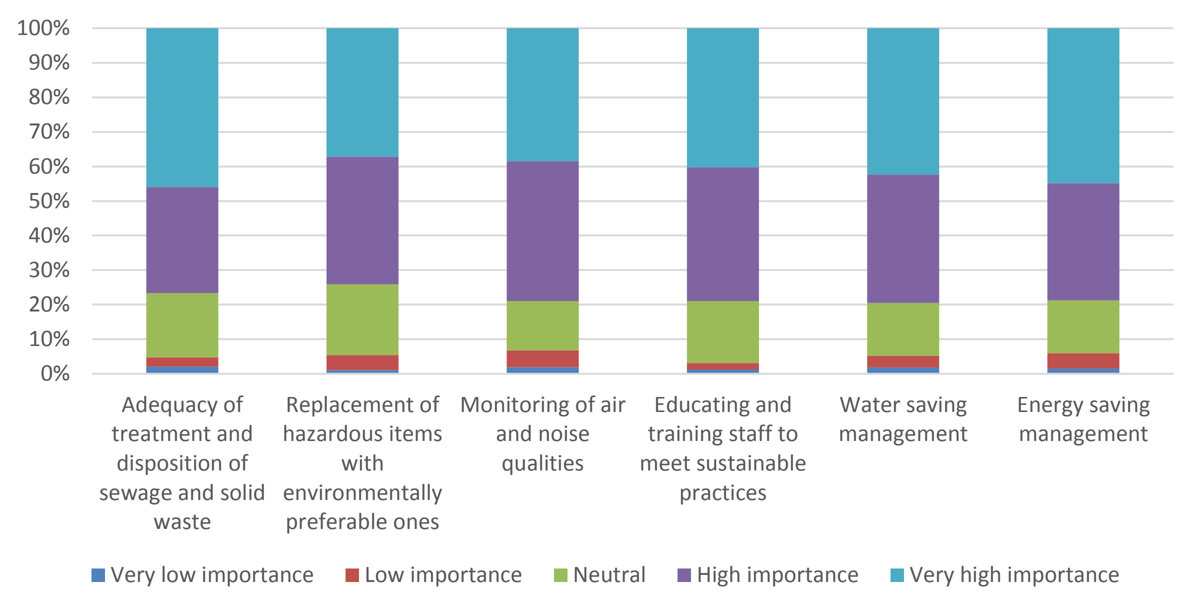
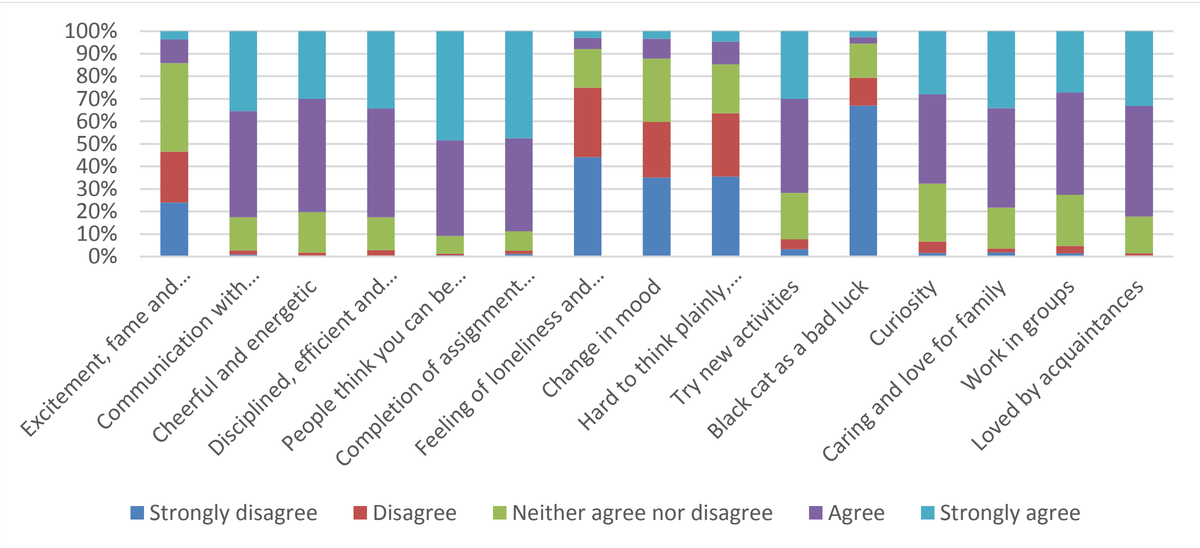



**FIGURE 14: PERSONALITY TRAITS**




## Hotel Quality

Around 51% of the tourists interviewed expressed utter satisfaction for their hotel accommodation compared to only 1.8% who strongly disagreed. 12.2% of the tourists were neither satisfied or unsatisfied by the hotel or accommodation quality.

## Casino Visits

Around 81.7% of the tourists interviewed expressed utter disagreement for the casino visits compared to only 2.9% who strongly agreed. 7.7% of the tourists neither approved nor disapproved of the casino visits.

**Sun and Sea**

Around 63.2% of the tourists interviewed completely approved of the sun and sea as a factor to travel to Mauritius compared to only 2.7% who strongly disagreed. 10.3% of the tourists neither agreed nor disagreed with the sun and sea.

**Security**

Around 34.9% of the tourists interviewed expressed strong agreement for the level of security being effective in Mauritius while only 1.5% strongly disagreed. 14.8% of the tourists showed neutrality with respect to the level of security.

**Gastronomy**

Around 12% of the tourists interviewed expressed strong agreement with the local gastronomy being a principal factor to travel to Mauritius compared to 9.7% of the respondents who strongly disagreed. 34.9% of the tourists demonstrated neutrality.

**Hospitability**

Around 55.9% of the tourists interviewed expressed utter satisfaction for their hotel accommodation compared to only 1.3% who strongly disagreed. 7.4% of the tourists were neither satisfied or unsatisfied by the hospitality.

**Nature and Climate**

Around 53.4% of the tourists interviewed expressed utter satisfaction for their hotel accommodation compared to only 2.3% who strongly disagreed. 11.6% of the tourists were neither satisfied or unsatisfied by the hotel or accommodation quality.

**Shopping**

Around 4.7% of the tourists interviewed expressed utter satisfaction for their hotel accommodation compared to only 35% who strongly disagreed. 28.1% of the tourists were neither satisfied or unsatisfied by the hotel or accommodation quality.

**Internet and Telecommunications**

Around 22.8% of the tourists interviewed expressed utter satisfaction for their hotel accommodation compared to only 10.8% who strongly disagreed. 26.1% of the tourists were neither satisfied or unsatisfied by the hotel or accommodation quality.

**Overall Satisfaction**

Around 53.4% of the tourists interviewed expressed utter satisfaction for their hotel accommodation compared to only 1.5% who strongly disagreed. 4.3% of the tourists were neither satisfied or unsatisfied by the hotel or accommodation quality.

**Tourism Economic Impact – Crowding-out Effect**

Around 6.3% of the tourists interviewed expressed utter satisfaction for their hotel accommodation compared to only 7.3% who strongly disagreed. 57.1% of the tourists were neither satisfied or unsatisfied by the hotel or accommodation quality

# CHAPTER 7: RESULTS AND DISCUSSIONS

Our objective is to evaluate the length of stay of the tourists who choose to holiday in Mauritius. Boxplots illustrating the average duration, by several explanatory terms are provided below from Figure 15 to Figure 19.

Before turning to more complicated, multivariate survival analysis, simple non-parametric models were executed. Stata 14.1 software was used to conduct out estimation process. We will start with the most classic of survival analysis that is the renowned Kaplan-Meier methodology. A variety of survival equality tests is also computed to validate any difference in the survival functions as per some predetermined consortiums. Subsequently the Nelson-Aalen hazard estimators are evaluated.

First, by following a graphical approach the Kaplan-Meier survival function is plotted for all variables. The number of tourists that leave Mauritius after spending a certain number of days can be used to obtain the Kaplan-Meier survival function for the whole population. At the beginning, since all the tourists are at staying state, the survival function has a value of 1. As time goes on, tourists leave Mauritius at random times. The last tourist has left Mauritius after staying 60 nights. The survival probabilities gives the chances that a tourist stays for a certain number of nights. For example, the probability that tourist stays in Mauritius for 9 nights is 48%. Such Kaplan-Meier estimations are plotted in Figure below.

There are no issues with data censoring in the duration model as the entire information throughout the whole period is available under analysis. The dependent variable of this study is the length of stay in Mauritius as measured in terms of days, of tourists, homeward bound from vacations.

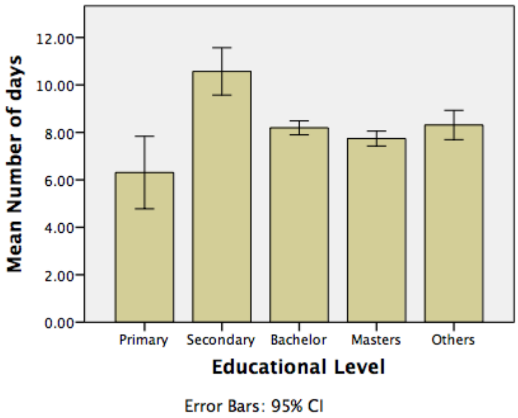



FIGURE 15: DURATION OF STAY, BY EDUCATIONAL LEVEL

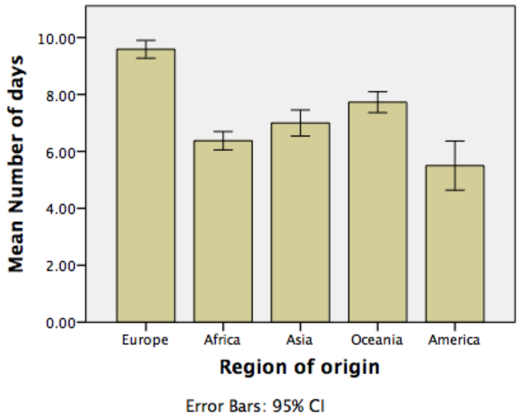



FIGURE 16: DURATION OF STAY, BY REGION OF ORIGIN

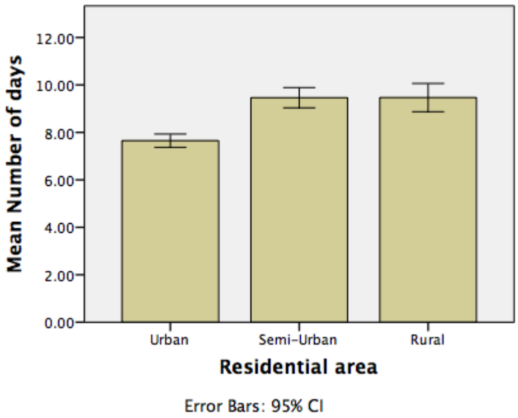



FIGURE 17: DURATION OF STAY, BY RESIDENTIAL AREA

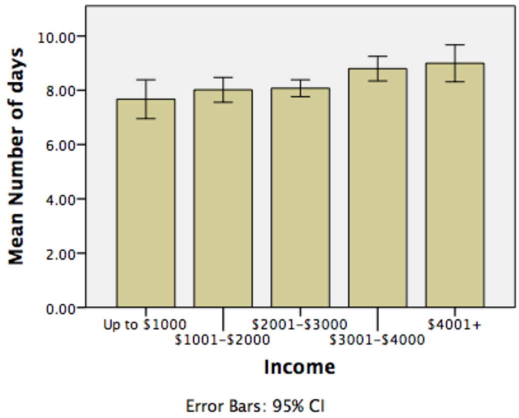



FIGURE 18: DURATION OF STAY, BY INCOME

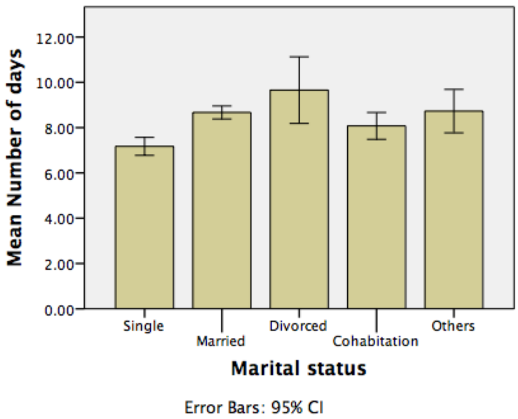
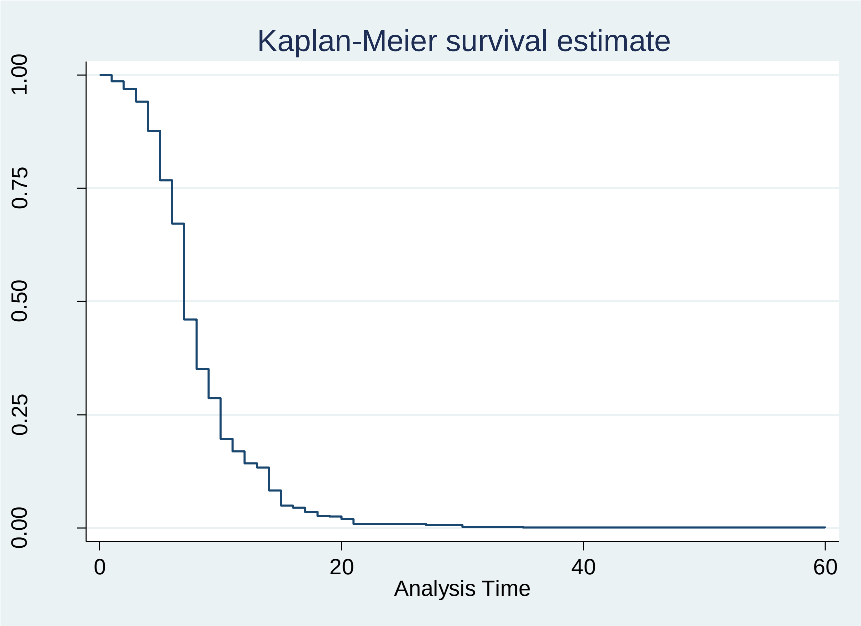



FIGURE 19: DURATION OF STAY, BY MARITAL STATUS




**SOURCE: COMPUTED**

Covariates used to explain the length of stay are categorised into different groups. Age is introduced in its linear form and monthly income is measured in US dollars. Fifty-one countries of origin from where the 1647 tourists came to Mauritius are identified by region-specific dummy variables.

The Cox regression model and Weibull model are then used to estimate the determinants of the length of stay in Mauritius. The results are presented below (For the Weibull, the results are in the appendix). They are not necessarily comparable as the Cox regression model is in proportional hazards form. In interpreting the values of beta in Cox regression, negative values suggest a negative effect on the hazard rate, therefore a positive effect on the survival rate.

The most important assumption of Cox's PH specification is that the hazard ratio is proportional over time. In survival analysis, the interpretation of results is distinctive from that of the conventional linear regression. It is noteworthy that the results of the Cox regression and the Weibull model are given in terms of their coefficients of the explanatory variables. A positive coefficient indicates that an increase in the explanatory variable has an increasing impact on the length of stay whilst a negative coefficient reports that as the value of a variable increases, the survival of the duration decreases.

In general, both Cox and Weibull's regressions yield similar results for most coefficients in the regression. There are four groups of regressors, for the sake of clarity of discussion. The first group pertains to the tourists' socio-demographic profiles while the second section characterises diverse trip attributes. The third group relates to tourists' reported importance of sustainability practices. The fourth and final group pertains to the five personality traits.

Those covariates which have a statistically significant effect on the survival probability functions are specified according to their estimated p-values.

**TABLE 7: COX'S REGRESSION**

| Variables | Coefficient | Standard Error |
|---|---|---|
| Age | -0.0078701 | 0.002181*** |
| **Gender (Baseline: Female)** | | |
| Male | 0.0728958 | 0.0593379 |
| **Educational Attainment (Baseline: Bachelor)** | | |
| Primary | 0.7124662 | 0.2906769** |
| Secondary | -0.1893652 | 0.0870885** |
| Masters | 0.130039 | 0.0628745** |
| Others | 0.0327948 | 0.0877817 |
| **Country Region (Baseline: Africa)** | | |
| Europe | -0.7204122 | 0.0801361*** |
| Asia | -0.3351398 | 0.1022509*** |
| Oceania | -0.5142924 | 0.2352148** |
| America | 0.1217378 | 0.2451047 |
| **Regional Location (Baseline: Rural)** | | |
| Urban | 0.1713819 | 0.0581799*** |
| **Income (Baseline: Income Level 1)** | | |
| Income 2 | -0.1664231 | 0.1052307 |
| Income 3 | -0.1423691 | 0.1071097 |
| Income 4 | -0.0432597 | 0.1157023 |
| Income 5 | -0.0261201 | 0.1193747 |
| **Marital Status (Baseline: Single)** | | |
| Married | -0.1530913 | 0.0744455** |
| Divorced | -0.2153308 | 0.141628 |
| Cohabitation | -0.0426789 | 0.1193 |
| Others | 0.078038 | 0.1632236 |
| **Travel Expenditure** | -0.0001719 | 0.0000135*** |
| **Purpose of Visit (Baseline: Others)** | | |
| Holiday | 0.0718439 | 0.2272516 |
| Business | -0.0281604 | 0.2649447 |
| Visit Friends and Relatives | -0.8309323 | 0.2557725*** |
| Conference | -0.606103 | 0.3847149 |
| Previous Visits | -0.0090461 | 0.0063451 |
| **Trip Attributes** | | |
| Hotel quality | 0.2306981 | 0.0408696*** |
| Casino Visits | -0.0436414 | 0.0326024 |
| Sun and Sea | -0.1797679 | 0.0375514*** |
| Security | 0.069428 | 0.0394783* |

| | | |
|---|---|---|
| Gastronomy | -0.0178125 | 0.0277042 |
| Hospitability | -0.0851677 | 0.0400198** |
| Nature and Climate | -0.0599437 | 0.0357406* |
| Shopping | 0.0099282 | 0.0272081 |
| Internet and Telecommunications | 0.0299885 | 0.0239305 |
| **Tourism Economic Impact** | | |
| Crowding-Out | -0.0148581 | 0.0325562 |
| **Environmental Initiatives** | | |
| Waste and Sewage Management | -0.0315784 | 0.0381601 |
| Quality Environmental Management | -0.0512078 | 0.0576556 |
| Water Saving Management | -0.0382113 | 0.0579924 |
| Energy Saving Management | 0.0685306 | 0.0574749 |
| **Personality Factors** | | |
| Extraversion | 0.0412012 | 0.0527993 |
| Conscientiousness | -0.0966657 | 0.0530689* |
| Neuroticism | -0.0370943 | 0.0336619 |
| Openness | 0.2286719 | 0.0490426*** |
| Agreeableness | 0.0620915 | 0.0518328 |

**SOURCE: COMPUTED.**

**ASTERISKS (***, **, *) DENOTE STATISTICAL SIGNIFICANCE RESPECTIVELY AT 1%, 5% AND 10% LEVEL. ALL MODELS WERE ESTIMATED IN STATA 14.**

None of income levels, casino visits, gastronomy, shopping, internet and telecommunications, crowding-out, environmental initiatives, extraversion, neuroticism and agreeableness appear to have any significant effect on the length of stay.

The category of reference in the general model is one individual who is a female, completed Bachelors degree, national from the African region, resident of rural area, earning a monthly income below $1000, single and coming to Mauritius for other purposes.

With respect to the first covariate, the relationship between the age variable and length of stay is negative and statistically significant, implying that the greater the age, the shorter the stay at the destination. This might suggest that young people take longer trips than people from older age groups on long-haul visits especially.

In general, country region has been found to play a significant role in determining the length of stay in Mauritius. Place of origin is a significant determinant of tourists' length of stay. From Table, belonging to the European region seems to reduce the probability of staying longer by 83%. Compared to Africans, the Europeans, Asians and Oceanians are the ones staying shorter. As for the income variable, it indicates that people who earn a monthly salary between USD

1000 and USD 2000 have less probability to stay longer compared to people earning less than USD 1000, which is taken as a base dummy.

Compared to single tourists, married tourists are inclined to stay shorter in Mauritius. Most of the time, these tourists may be travelling with their spouses and/or kids, implying that they spend on their family. Ganzon and Fillone (2014) suggest that they may not have the luxury and freedom of spending more time on vacations as compared to single free-travellers. It will then be advantageous to select single tourists as the target market to promote longer stays in Mauritius.

Travel expenditure includes spending on various items such as accommodation, transport, food and beverages, shopping and entertainment. It was found to have a negative impact on the length of stay as reported by various studies. This result is in line with Gokovali. A required large budget increase the hazard of ending the tourists' trip as tourists have limited travel budget. An inspection into the aspects such as transportation, accommodation, food and entertainment of their vacation signals a high budget allocation. Expenditure has a negative impact on hazards of length of stay which is rational for tourists with budgetary constraints.

The results show that the trips carried out by international tourists in order to visit family members and friends have a significantly stronger impact on length of stay than the other proposed reasons for travelling (holiday, business, conference). The shorter duration of this type of trip with respect to "others" category could be related. Those who frequently visited Mauritius in the past are associated with a shorter stay. This is in contrast with the findings of Gokovali which stipulate that as the number of previous visits to the destination increases, the probability of staying longer also increases. A frequent visitor to Mauritius would allocate holiday time to various destinations compelling him to spend as little time as possible at a particular destination, as reported by Kazururu.

Casino visits in Mauritius were found to be significant and have shorter impact on the length of stay in Mauritius. The presence of having casino visits does not help to improve the length of stay. The sun and sea factor were found to decrease the probability of staying longer in Mauritius. The sign of hospitability is somewhat paradoxical because Mauritius is a well-known destination for its heart-warming nature towards tourists. This might imply that tourists are reluctant to pay attention to this covariate while making decision on how long to stay. Further results demonstrate that being vigilant, efficient and having a sense of organisation reduces the probability of staying longer.

Estimations show that being a male tourist increases the length of stay with respect to the category female. Further investigations claim educational attainment to be a significant covariate contributing to longer stays. Relative to the Bachelor treated as dummy, tourists with the lowest levels of education (Primary) have higher survival rates than those with secondary education whilst no significant differences have been detected among others which included mainly PhD holders and professionals.

Further results unfurl that being a tourist from the urban location increases the length of stay. Hotel quality is another vital issue in this research, given that the aim was to focus on the differences in stay determinants. Tourists satisfied with their hotel accommodation were found to stay longer. Differences seem to be significant.  The level of security established in Mauritius increases the probability of staying longer. The digital age sparks the fact internet and telecommunications are found to be significant in increasing the length of stay.

Having an open attitude to life increases the probability of staying longer.  Being of agreeable nature has proved to be significant in increasing the length of stay in Mauritius.

# CHAPTER 8: CONCLUSION AND POLICY IMPLICATIONS

Worldwide tourism demand is evolving swiftly towards a soar in the number of annual trips. This report has highlighted the determinants of tourists' length of stay in Mauritius. Among the key determinants are the tourists' socio-demographic characteristics, trip-related attributes and personality traits which appear to be the most influential. Therefore these summon the most attention by the stakeholders in the industry.

One limitation of this study is that as the questionnaire was self-administered, the accuracy of the responses would hinge upon the memory, recall and honesty of the tourists.

The development of tourism sector is as inevitable economic strategy. The strategies in the tourism sector that would propel Mauritius forward are: during low-season, honeymooners should be the main target. The operational and tactical side of tourism considers that the segments should be filled in during both low and high seasons. In retrospection, 2008-2010 can be labelled as the years of consolidation and investment in the future of the product. How do we sustain the industry now?

We are shifting from the traditional markets. For example, UK has a risk of not performing in the future. South Africa, which is a key market, went into a recession. The point is that the main traditional markets might now yield the same amount of tourists in the future. Per se, we are already analysing the various sides of the box by promoting Mauritius as a honeymoon and family destination to Chinese and Asian travellers.

Single tourists are inclined to stay longer in Mauritius. Most of the time, these tourists may be travelling with their friends or co-workers, implying that they only spend for themselves. Their travel group may also be able to engage in more and specific activities as they may most likely be of the same age and similar interests. Also, since they are single, they may have the luxury and freedom of spending more time on vacations as compared to married tourists. It will then be advantageous to select single tourists as the target market in promoting a longer stay in the province. Tourism creates opportunities for a plethora of players such as handicraft-makers, technologists, hoteliers, to name a few. Tourism can be the future of sustainable development.

Most importantly, marketing strategies should address personality traits so that Mauritius becomes the top-of-the-mind tourist destination for travellers.

## References


- Alegre, J., and Pou, L. (2006). The length of stay in the demand for tourism. *Tourism Management, 27, 1343–1355.*
- Artal Tur, A., García Sánchez, A. and Sánchez García, J.F. (2008) The length of stay determinants for sun-and-sand tourism: An application for the Region of Murcia, *XVI Jornadas ASEPUMA – IV Encuentro Internacional*
- Barros, C, P., Butler, R. and Correia, A. (2010). The length of stay of golf tourism: A survival analysis, *Tourism Management, 3, 13–21.*
- Barros, C. P., Correia, A. and Crouch, G. (2008) Determinants of the length of stay in Latin American tourism destinations, *Tourism Analysis, 13, 329–340.*
- Boyce, C. J. and Wood, A. M. (2011). Personality and the marginal utility of income: Personality interacts with increases in household income to determine life satisfaction. *Journal of Economic Behavior Organization, 78, 183– 191.*
- Bresson, G. and Logossah, K., (2008). Crowding-out effects of cruise tourism on stay-over tourism within the Carribean. A non-parametric panel data evidence
- Candela, G. and Figini, P., (2012). The Economics of Tourist Destinations. *Springer Science and Business Media*
- Cernat, L. and Gourdon, J. (2007) Is the concept of sustainable tourism sustainable? Developing the sustainable tourism benchmarking tool. *United Nations Conference on Trade and Development.*
- Chaiboonsri, C, and Chaitip, P. (2012) The Modeling International Tourism Demand for Length of Stay in India: Social Development, Economic Development, Journal of Knowledge Management, Economics and Information Technology, 2 (2), 1-17.
- Cho, H.-S., Byun, B. and Shin, S., (2014). An examination of the relationship between rural tourists' satisfaction, revisitation and information preferences: A Korean Case Study. *Sustainability, 6, 6293-6311, ISSN 2071-1050.*
- Cox, D. R. (1972) Regression Models and Life Tables (with Discussion), *Journal of the Royal Statistical Society, Series B, 34, 187—220*.
- Craggs, R. and Schofield, P., (2009). Expenditure-based segmentation and visitor profiling at the Quays in Salford, Uk, *Tourism Economics, Vol 15(1), 243-260.*
- Crompton, J., (1979). Motivations of pleasure vacations. *Annals of Tourism Research, 6(4): 408-424*

- Cuddeback, G., Wilson, E., Orme, J. G., and Combs-Orme, T. (2004), Detecting and Statistically Correcting Sample Selection Bias, *Journal of Social Service Research, 30 (3), 19-33.*
- De Menezes, A.G., Moniz, A. and Viera, J.C. (2008) The determinants of length of stay of tourists in the Azores, *Tourism Economics, 14 (1), 205–222.*
- De Menezes, A.G., Vieira, J.C. and Carvalho, J. (2009). Assessing tourist satisfaction in the Azores: A microeconometric approach. European Journal of Tourism Research, Vol. 2, No.1
- De Oliveira Santos, G, E., Ramos, V., and Rey-Maquieira, J. (2014), Length of Stay at Multiple Destinations of Tourism Trips in Brazil, *Journal of Travel research, 1-13.*
- Decrop, A. and Snelders, D., (2004). Planning the summer vacation: An Adaptable Process. *Annals of Tourism Research, Vol. 31, No. 4, pp. 1008–1030*
- Disegna, M. and Osti, L., (2016). Tourists' expenditure behaviour: the influence of satisfaction and the dependence of spending categories. *Tourism economic, 22(1), 5-30*
- Eysenck, H. J. (1959). Manual of the Maudsley Personality Inventory. London: University of London Press.
- Fleischer, A.Y. and Pizam, A. (2002). Tourism Constraints among Israeli Seniors. *Annals of Tourism Research, 29, 1, pp. 106-123.*
- Gabbi, G. and Zanotti, G., (2011). Emotional state financial expectations and overconfidence. *Siena Working Paper in Finance. Retrieved from http://efmaefm.org/ 0EFMSYMPOSIUM/ Renmin-2011/papers/Gabbi.pdf*
- George, B. (2002). The relationship between lottery ticket and scratch-card buying behaviour, personality and other compulsive behaviours. *Journal of Consumer Behaviour, 2, 7–22. doi:10.1002/cb.86.*
- Gokovali, U., Bahar, O., and Kozak, M. (2007) Determinants of length of stay: A practical use of survival analysis, *Tourism Management, 28 (3), 736–746.*
- Goldberg, L. R. (1993). The structure of phenotypic personality traits, *American Psychologist, 48, 26–34.*
- Han, H., Hsu, L.J. and Lee, J. (2009). Empirical investigation of the roles of attitudes toward green behaviours, overall image, gender, and age in hotel

customers' eco-friendly decision-making process, *International Journal of Hospitality Management, 28, 519-528.*

- Heckman, J. J. (2001). Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture. *Journal of Political Economy, 109(4), 673–748.*
- Jaunky V. C. and Ramchurn, B. (2014) Consumer behaviour in the scratch card market: A double-hurdle approach, International Gambling Studies 14, 1, pp. 96-114.
- Kazururu, B., (2014). Determinants of Tourist Length of Stay in Tanzania. *International Journal of Business and Social Science, Vol. 5, No. 9(1)*
- Kozak, M. (2001) Repeaters' behaviour at two distinct destinations, *Annals of Tourism Research 28(3), 784-807.*
- Lancaster T. (1990) The Econometric Analysis of Transition Data, Cambridge, *Cambridge University Press.*
- Lancaster, K. J., (1966). A new approach to consumer theory. *Journal of Political Economy, 74, 132–157.*
- Logossah, K., Maupertui, M.-A,. (2007). La spécialisation touristique des petites économies insulaires en développement est-elle une voie de croissance durable? *Vol 5(1), 35-55.*
- Machado, L, P. (2010) Does destination image influence the length of stay in a tourism destination? *Tourism Economics, 16 (2), 443–456.*
- Martínez-Garcia, E. and Raya, J. M. (2008) Length of stay for low-cost tourism', *Tourism Management, 29, 1064–1075.*
- Masau, P. and Prideaux, B., (2003). Sustainable tourism: A role for Kenya's hotel industry. Current Issues in Tourism, 6:3, 197-208, DOI: 10.1080/13683500308667953.
- McCool, S. F., & Martin, S. R. (1994). Community attachment and attitudes toward tour-ism development. *Journal of Travel Research,32(3), 29–34.*
- McCrae, R. R., and Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers, *Journal of Personality and Social Psychology, 52, 81-90.*
- Menezes, A. G., Moniz, A., and Vieira, J. C. (2008) The determinants of length of stay of tourists in the Azores, *Tourism Economics, 14, 205–222.*
- Menezes, A., A. Moniz, Vieira, J. (2008) Determinants of length of stay: A parametric survival analysis. *Tourism Economics, 14(1), pp. 205-222.*

- Mensah, I. and Mensah, R.D., (2013). International Tourists' Environmental Attitude towards Hotels in Accra. *International Journal of Academic Research in Business and Social Sciences, Vol. 3, No. 5, ISSN:2222-6990*
- Meyer, B. (1990) Unemployment Insurance and Unemployment Spells, *Econometrica, 58(4), 757-782.*
- Mowen and Harris EG (2003) The MDPS Method of message theme development: a new tool for managers, *Journal of Consumer Marketing, 20(12), 1045-1066.*
- Mowen, J. C. (2000). The 3M Model of Motivation and Personality: Theory and Empirical Applications to Consumer Behavior. Boston, MA: Kluwer Academic.
- Peypoch, N., Randriamboarison, R., Rasoamananjara, F. and Solonandrasana, B. (2012) The length of stay of tourists in Madagascar, *Tourism Management, 33, 1230-1235.*
- Prentice, R. L. and Gloeckler, L. A. (1978) Regression analysis of grouped survival data with application to breast cancer data, *Biometrics, 34, 57–67*.
- Saarinen, J. (2006) Traditions of sustainability in tourism studies, *Annals of Tourism Research*, Vol 33, No 4, pp 1121–1140.
- Salman, A. and Hasim, M. S., (2012). Factors and Competitiveness of Malaysia as a tourist destination: A study of outbound middle east tourists. *Asian Social Science, Vol. 8, No. 12*
- Scheyvens, R. and Momsen, J.H. (2008). Tourism and Poverty Reduction: Issues for Small Island States. Tourism Geographies, 10:1, 22-41, DOI: 10.1080/14616680701825115.
- Schneider, P.P. and Vogt, C.A., (2012). Applying the 3M Model of Personality and Motivation to Adventure Travelers. *Journal of Travel Research, 51(6) 704-716*
- Schultz, D. P., and Schultz, S. E. (2008). Theories of personality (9nd ed.). *Wadsworth: Cengage Learning.*
- Su, Y.-W., Lin, H.-L. and Liu, L.-M., (2012). Chinese tourists in Taiwan: Crowding out effects, opening policy and its implications. *Tourism Management Perspectives 4 45-55*
- Tran, X., Nguyen, B. L. and Nguyen, M.C. (2015) Effects of the Big Five Personality Traits on Recreation Types - The Case of Vietnam Tourism. *Tourism Travel and Research Association: Advancing*

- Tussyadiah, I.P. and Pesonen, J., (2015). Impacts of peer-to-peer accommodation use on travel patterns. *Journal of Travel Research 1-19*
- UNCED (1992), Agenda 21: A Guide to the United Nations Conference on Environment and Development. *UN Publications Service, Geneva.*
- Varian, H. R. (1987). Intermediate microeconomics: A modern approach. *New York:W.W. Norton & Co.*


- Varian, H. R. (1987). Intermediate microeconomics: A modern approach. *New York: W.W. Norton & Co.*

**COMMENTS FROM SURVEY RESPONDENTS**

(1) Too expensive as a tourist destination. No value for money. German travellers would prefer to go to Mexico. Food and beverages overpriced in the departure lounge – a great business model but not worth it.

(2) Transport costs occupy a disproportionate share of the cake. Taxi drivers won't make concessional prices to attract tourists = overestimation of the actual price; especially prices are hiked three to four times for European tourists, assuming they should be paid in euros.

(3) Bleak customer service. Mauritian shops should be more professional in selling their products. Lack of professionalism and efficiency. Criticised as not very courteous.

(4) The roads in Mauritius are not very safe and they do not guide very well. A lack of signage - dreadful. Careless and rash drivers on the road. Tourism in Mauritius comes at a cost. The lives are endangered.

(5) Sustainability issue? The beaches are cluttered with plastic bottles, thrown take-aways, plastic bags. Environmental awareness heading south. Loyal French visitor thinks that she should seldom come to Mauritius now. Measures takes to prevent beach environmental degradation are barely enough – disparaging and discouraging.

(6) Recreational activities/leisure

The main tourist attraction sites can be visited in 5-6 days. No motivation or incentive to stay longer compared to Sentosa Island.

No kids' corner in the departure lounge. No magazines or newspapers to kill time during transits in Mauritius. No smoking area in the lounge. No foreign currency exchange dealer point after security checks.

(7) Lack of a night life culture (most shops are closed by 6 pm).

(8) Promoted as a paradise island for international holiday-makers but a veritable hell on earth for dogs.

# Appendix

**TABLE 8: WEIBULL'S REGRESSION**

| Variables | Coefficient | Standard Error |
|---|---|---|
| Constant | -4.474072 | 0.4136811*** |
| Age | -0.0092742 | 0.0021776*** |
| **Gender (Baseline: Female)** | | |
| Male | 0.1485197 | 0.059656** |
| **Educational Attainment (Baseline: Bachelor)** | | |
| Primary | 0.7645388 | 0.2914233*** |
| Secondary | -0.2954325 | 0.0887281** |
| Masters | 0.1818844 | 0.0633538*** |
| Others | 0.036143 | 0.0886836 |
| **Country Region (Baseline: Africa)** | | |
| Europe | -0.8307273 | 0.0812825*** |
| Asia | -0.4568271 | 0.1040896*** |
| Oceania | -0.4300526 | 0.2365279* |
| America | 0.1466136 | 0.2393395 |
| **Regional Location (Baseline: Rural)** | | |
| Urban | 0.1411676 | 0.0584434** |
| **Income (Baseline: Income Level 1)** | | |
| Income 2 | -0.2617266 | 0.105574** |
| Income 3 | -0.2232543 | 0.1073682** |
| Income 4 | -0.1027534 | 0.1161453 |
| Income 5 | -0.1662634 | 0.1200177 |
| **Marital Status (Baseline: Single)** | | |
| Married | -0.1709093 | 0.074423** |
| Divorced | -0.1814675 | 0.1416055 |
| Cohabitation | 0.0427492 | 0.119291 |
| Others | 0.2338341 | 0.1631733 |
| **Travel Expenditure** | -0.0002066 | 0.0000136*** |
| **Purpose of Visit (Baseline: Others)** | | |
| Holiday | 0.0245622 | 0.2263627 |
| Business | -0.2219564 | 0.2658796 |
| Visit Friends and Relatives | -1.484955 | 0.2575345*** |
| Conference | -0.9082978 | 0.3829073** |
| **Previous Visits** | -0.0110806 | 0.0065451* |
| **Trip Attributes** | | |
| Hotel quality | 0.2766522 | 0.0422544*** |
| Casino Visits | -0.0977876 | 0.0331058*** |
| Sun and Sea | -0.2189934 | 0.036856*** |
| Security | 0.0895322 | 0.0396814** |
| Gastronomy | -0.0196528 | 0.0283019 |
| Hospitability | -0.1398942 | 0.0407858*** |
| Nature and Climate | -0.0136621 | 0.0363617 |

| | | |
|---|---|---|
| Shopping | 0.0031079 | 0.0268901 |
| Internet and Telecommunications | 0.0418505 | 0.0241966* |
| **Tourism Economic Impact** | | |
| Crowding-Out | -0.0106208 | 0.0336804 |
| **Environmental Initiatives** | | |
| Waste and Sewage Management | -0.0337511 | 0.0384363 |
| Quality Environmental Management | -0.0762231 | 0.0573064 |
| Water Saving Management | -0.0219383 | 0.0580883 |
| Energy Saving Management | 0.0410277 | 0.057091 |
| **Personality Factors** | | |
| Extraversion | 0.0437596 | 0.0523704 |
| Conscientiousness | -0.1456167 | 0.0539334*** |
| Neuroticism | -0.0355219 | 0.0333898 |
| Openness | 0.2640318 | 0.0498201*** |
| Agreeableness | 0.1041091 | 0.0526446** |
| | | |
| /ln_p | 1.049629 | 0.018329 |
| | | |
| P | 2.856593 | 0.0523586 |
| 1/p | 0.3500674 | 0.0064164 |

**SOURCE: COMPUTED.**

Asterisks (***, **, *) denote statistical significance respectively at 1%, 5% and 10% level. All models were estimated in Stata 14.