ISSN: 2319-9490

*Available online at http://www.ijcrls.com*

*International Journal of Current Research in Life Sciences*
*Vol. 07, No. 04, pp.1603-1616, April, 2018*

# RESEARCH ARTICLE

# EXPLORING SEASONALITY PATTERNS FOR PASSENGER DATA IN SAUDI ARABIAN AIRLINES

**[1]Said Ali ElQuliti and [2,]*Abdulaziz Tariq Qadi**

[1]Professor, Department of Operations Research and Decision Support, Faculty of Computers and Information, Cairo University, Egypt
[2]HR Consultant, Saudi Arabian Airlines, Jeddah, Kingdom of Saudi Arabia

Received 25th February, 2018; Accepted 07th March, 2018; Published 15th April, 2018

## ABSTRACT

Forecasting of air traffic means to estimate the number of prospective passengers that use air transport. Time Series data for the number of passengers in Saudi Arabian Airlines are collected in Gregorian and partially in Hijri calendars. The data are recorded monthly for ten years from January 2007G to July 2017G in Gregorian. However, Hijri data were only from Muharram 1435H to Shawal 1438H. The Gregorian data are transformed to Hijri to complete the missing gab. A major question to be studied is raised about existence of seasonality and which data is better than the other in expressing seasonality pattern. Another important question is to identify months of the year having better seasonality than others. To explore the seasonality patterns for the time series data in Gregorian and Hijri calendars we used the following methods: Scatter Diagrams, Autocorrelation Functions, $x^2$Goodness-of-Fit Test, Seasonality Indexes using Ratio to Moving Average, Plot of Changes in the Seasonal Pattern and Autoregression to assess strength of seasonality. New methods are proposed to assess the seasonality of the data and moreover to identify which months of the year have better seasonality, these methods are: Month's Orders, Measures of Dispersion for Seasonality Indexes (Range, Quartile Deviation, Average Deviation and Standard Deviation).

***Key words:*** Seasonality Patterns, Test of seasonality, Time series, Saudi Airlines passengers.



**Citation: Said Ali ElQuliti and Abdulaziz Tariq Qadi, 2018.** "Exploring Seasonality Patterns for Passenger Data in Saudi Arabian Airlines" *International Journal of Current Research in Life Sciences*, 7, (04), 1603-1616.

## INTRODUCTION

Civil Aviation, through a complicated interaction with other econometric sectors, benefits from and contributes to the economic development of all nations. As incomes and production level increase, the demand for aviation services expands. Therefore prospective tourism, trade and employment could be forecasted as well. Civil Aviation is an important instrument in economic development, and air transport also provides an intangible benefits by facilitating the international treaties and understanding. On the other, the role of air transport as a catalyst for general economic and social development is due to the expedition and flexibility, which has been provided by the global air transport system. Forecasting is a prediction of future after studying the past. It is necessary especially nowadays in order to meet the anticipated demands. It is a planning tool that helps management for budgeting, planning, and estimating future growth. Recognizing and understanding the pattern of the data is the way to select appropriate forecasting technique for the given time series data. Planning for the future is one of the most important keys to success, forecasting is the way.

***Corresponding author:* Abdulaziz Tariq Qadi,**
Saudi Arabian Airlines, Jeddah, Kingdom of Saudi Arabia.

Forecasting studies and analyses the present and past data in order to predict the future, so one of the most important step in order to do an excellent in forecasting is recognizing and understanding the pattern of the data. A time series is a sequence of observations on a variable measured at successive points in time or over successive periods of time. The measurements may be taken every hour, day, week, month, or year, or at any other regular interval. The pattern of the data is an important factor in understanding how the time series has behaved in the past. If such behavior can be expected to continue in the future, we can use the past pattern to guide us in selecting an appropriate forecasting method. To identify the underlying pattern in the data, a useful first step is to construct a time series plot. A time series plot is a graphical presentation of the relationship between time and the time series variable; time is on the horizontal axis and the time series values are shown on the vertical axis. Let us review some of the common types of data patterns that can be identified when examining a time series plot. The underlying pattern in the time series is an important factor in selecting a forecasting method. Thus, a time series plot should be one of the first things developed when trying to determine what forecasting method to use. If we see a horizontal pattern, then we need to select a method appropriate for this type of pattern. Similarly, if we observe a trend and or seasonality in the data, then we need to use a forecasting method that has the capability to handle trend and or seasonality effectively. Understanding the nature of the data will help in discover which forecasting technique is more appropriate to the data. Time series data consists of observations on a set of variables over time. There are four data patterns which are horizontal (stationary), trend, cyclical and seasonal. A Stationary Series means that the mean and variance remain constant over time. Trend occurs when the long-term component that represents the growth or decline in the time series. The wavelike fluctuation around the trend is representing the cyclical pattern whereas a pattern of change that repeats itself year after year is representing the seasonal pattern.

## Background

Analysis of time series data for forecasting purposes involves the identification of patterns that exist in the data. These patterns may include: trend, seasonality, cyclicality and random variability. For studying seasonality, several statistical methods are available ranging from simple graphical techniques like drawing the scatter diagrams to more advanced statistical methods. Additionally, autocorrelation functions can be examined to assess regularity of periodicity or seasonality. Jones *et al.,* (1988) developed a test for determining whether incidence data for two or more groups have the same seasonal pattern. The $x^2$ goodness-of-fit test is relatively popular for detecting seasonality because of its simple mathematical theory, which makes it easy to calculate and understand (2016). Further, Marrero (1983) compared the performance of several tests for seasonality by simulation, which can be used as a guideline for selecting appropriate tests for a given data set based on the size of the data set and the shape of the sinusoidal curve. To apply any of these tests, however, observations must be aggregated into 12 monthly data points. Each test provides an indication of the presence or absence of statistical significance of seasonality, however, they do not provide a sense of the magnitude of seasonality or how much variance is explained by seasonal occurrence in the data. To test for significance of the seasonal pattern and to identify seasonality in specific months, Bquanttrading calculated the observed changes in a time series based on the month of a year (the difference between number of passengers in a month and the previous month) and then trace the pattern of changes for each month (Bquanttrading, 2015). Moineddin *et al.* (2003) propose the coefficient of determination of the autoregressive regression model fitted to the data as a measure for quantifying the strength of the seasonality. The performance of the proposed statistic is assessed through a simulation study and using two data sets known to demonstrate statistically significant seasonality.

## Data collection and inversion

It is required to gather the monthly data for passenger numbers in the last ten years; the data in the system are only available in Gregorian dates due to the nature of original inputs and the system setup (UBC Real Estate Division, 2011). As an exception case, passenger numbers are provided in Hijri dates only from Muharram 1435H to Shawal 1438H, whereas for Gregorian dates the data were given from January 2007G to July 2017G based on the request (Department of Network Planning and Scheduling). To convert the Hijri data into Gregorian for the monthly number of passengers, the following procedure is followed for the Hijri data from Muharram 1428H to Dhu AL-Hijjah1434H:

1. Monthly Gregorian data are converted to average daily number of passengers,
2. The start and end date for each Hijri month in every year are indicated and converted to Gregorian (International Astronomical Center webpage, 2018).
3. Each Hijri month is converted to the corresponding periods in Gregorian dates,
4. The total number of passengers in each Hijri month is calculated as the sum of daily average number of passengers in the corresponding periods in Gregorian dates.

As examples, Table 1 shows the number of passengers and the daily average for April and May 2008 G. and Table 2 shows the conversion of days for the month Rabi'll 1429 into the corresponding days in Gregorian and the total number of passengers in that month.

**Table 1. Number of passengers and the daily average for April and May 2008 G**

| Month | Total Monthly Number of passengers | Average daily number |
|---|---|---|
| April 2008 | 1,267,583 | 42253 |
| May 2008 | 1,296,791 | 41832 |

**Table 2. Calculating the total number of passengers in Rabi’II 1429**

| Days in Hijric 1429 | Days in Gregorian 2008 | Average number of passengers/day |
|---|---|---|
| Rabi’II 1- 24 | 7 April 2008 – 30 April | 42253 |
| Rabi’II 25 - 29 | 1 May – 5 May | 41832 |
| Total number of passengers in Rabi’II 1429 | | 1,223,226 |

Figures 1 and 2 show the scatter plots for the number of passengers in Gregorian and Hijri calendars. Both figures indicate an upward trend and an apparent seasonality.

## Exploring time series data patterns

There are many objectives for time series analysis:

- **Description:** Trend, seasonality/cyclicity, outliers, sudden changes or breaks.
- **Explanation:** Using one time series to explain another – May help understand the mechanisms.
- **Prediction:** Forecasting.
- **Control:** Time series often collected to improve control over a physical process and monitoring to alert when conditions exceed a priori determined threshold.

Time series data may have the following patterns: Random, Horizontal (stationary), Trend, Cyclical, and Seasonal. If a data set is random; then the successive values are not related to each other and are probabilistically independent of one another. For a random data almost all the autocorrelation coefficients are significantly equal to zero (Robert, 2017). A time series data is called horizontal (stationary) series when its mean and variance remain constant over time, if the series is stationary the Autocorrelation Function (ACF) decays fast to zero. A stationary time series has a mean, variance, and autocorrelation function that are essentially constant through time. The data is non-stationary when there is a large spike at lag 1 that slowly decreases over several lags (David Anderson *et al.* 2008). The trend is a long-term component that represents the growth or decline in the time series (Rao Vallabhaneni, 2017).

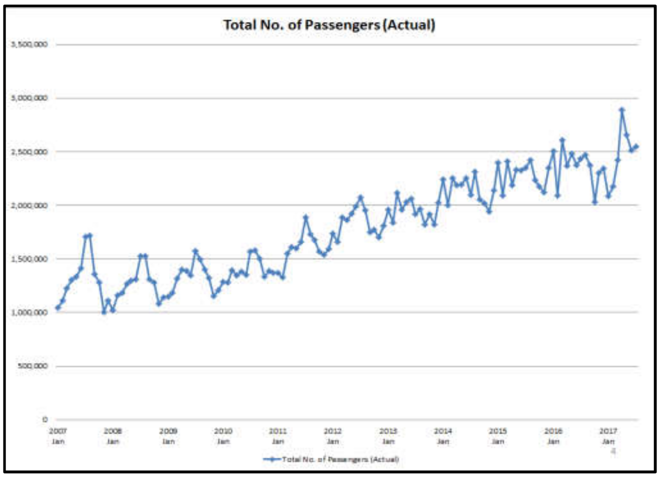



**Figure 1. Scatter plots for the number of passengers in Gregorian calendar**

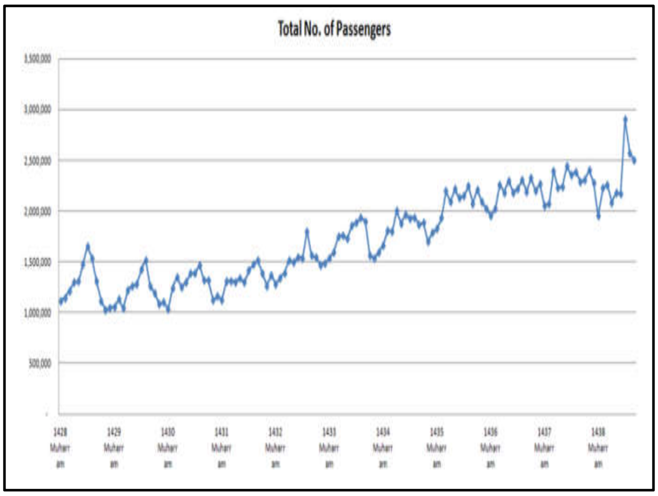



**Figure 2. Scatter plot for the number of passengers in Hijri calendar**

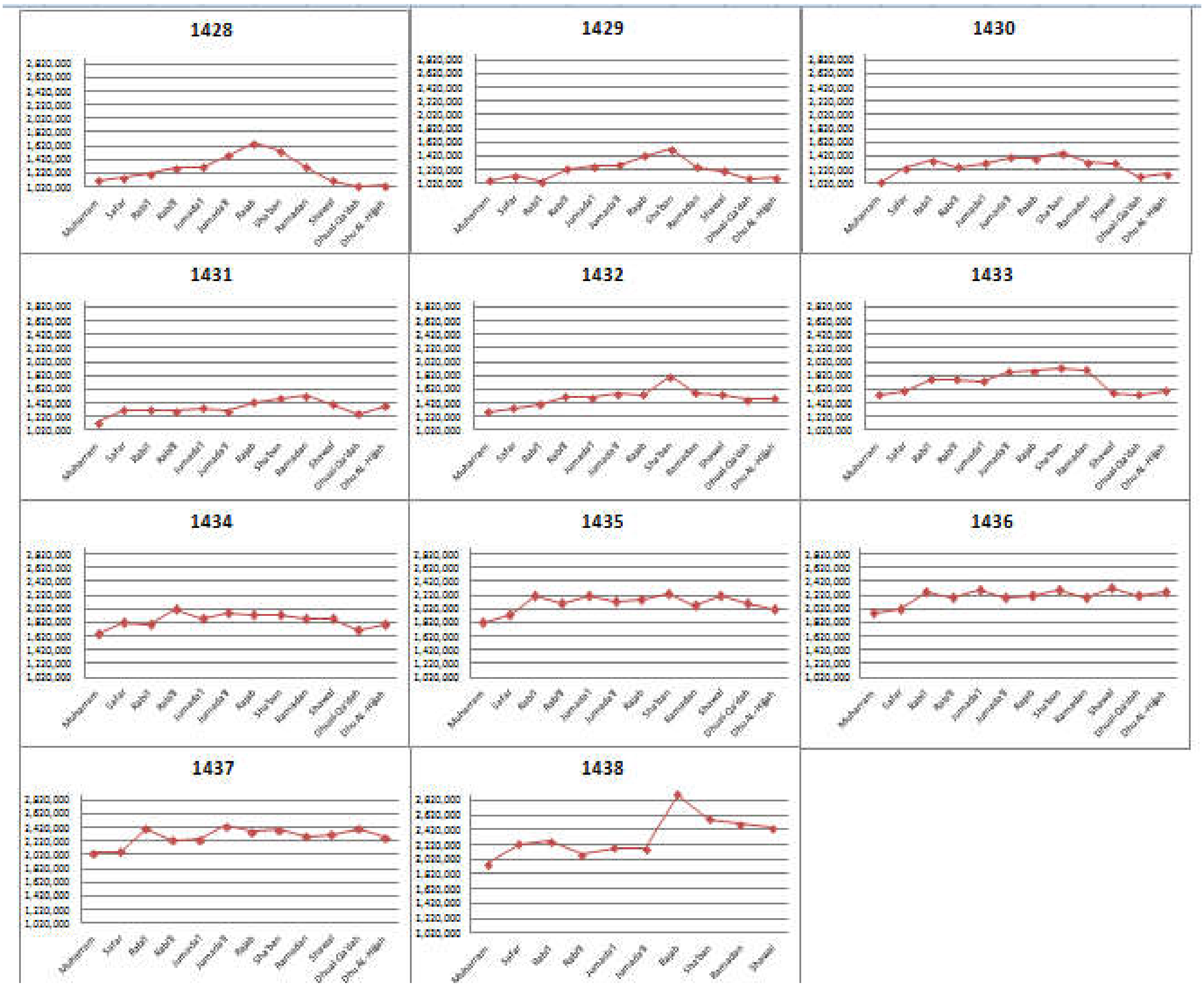
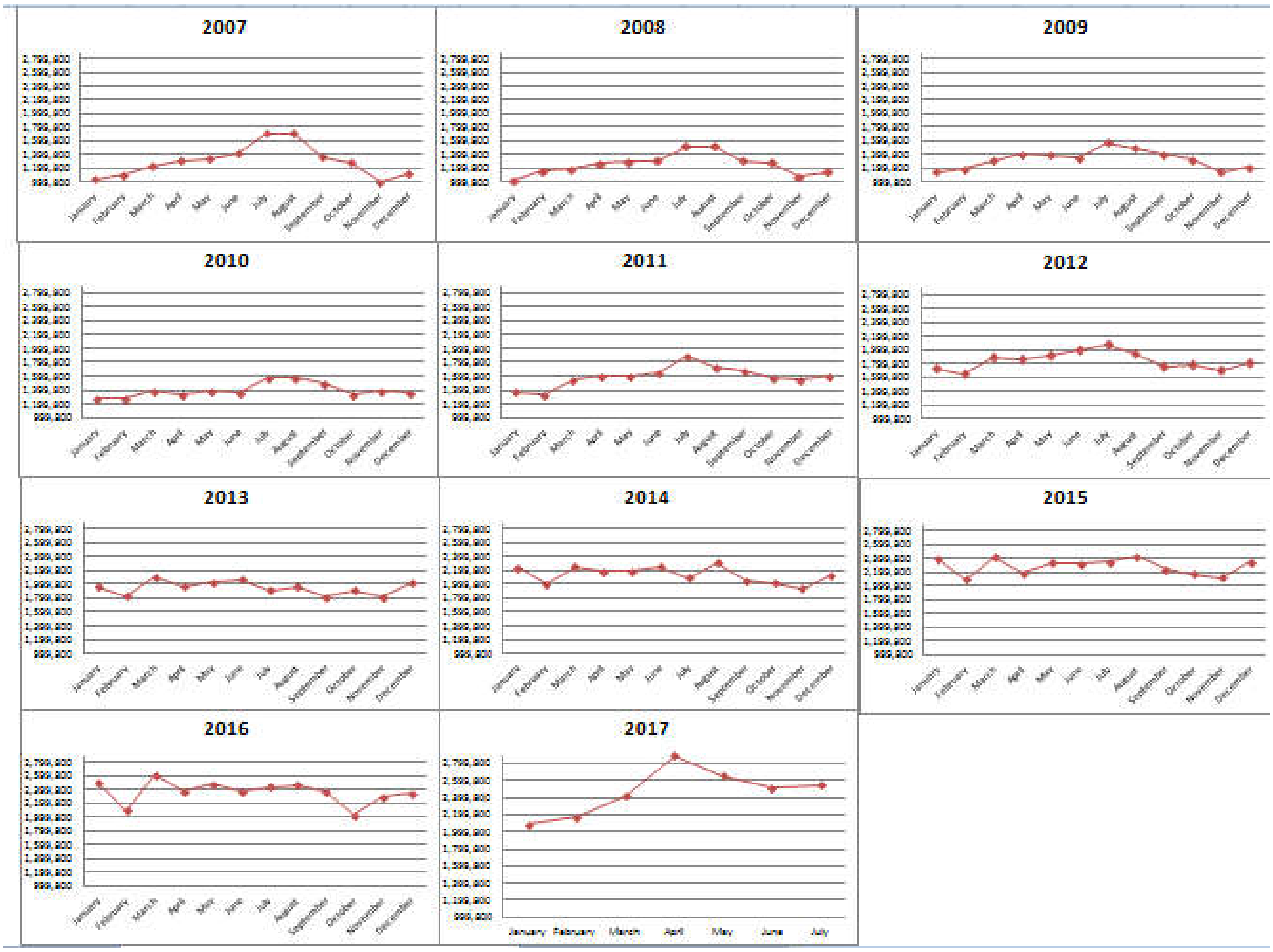



**Figure 3. Scatter diagram for passenger in Hijri data**

The cyclical component is a wavelike fluctuation around the trend usually affected by general economic conditions. Cyclical patterns tend to repeat in the data roughly every two, three, or more years. Cyclical fluctuations are usually influenced by changes in economic expansions and contractions, commonly referred to as the business cycle (Rob J Hyndman and George Athana¬sopou¬los 2016).

A cyclic pattern exists when data exhibit rises and falls that are not of fixed period. The duration of these fluctuations is usually of at least 2 years. Think of business cycles which usually last several years, but where the length of the current cycle is unknown beforehand. A seasonal pattern exists when a series is influenced by seasonal factors (e.g., the quarter of the year, the month, or day of the week). Seasonality is always of a fixed and known period. Hence, seasonal time series are sometimes called periodic time series. The seasonal component is a pattern of change that repeats itself year after year (Jain, 2007). Cyclical variations don't repeat themselves in a regular pattern, but they are not random variations either. Cyclical patterns are recognizable, but they almost always vary in intensity (the height from peak to trough) and timing (frequency with which the peaks and troughs occur). Since they cannot be accurately predicted, they are often analyzed with the irregular components.

policy, particularly with respect to the planning for seasonal demands for passenger services. Available statistical tests for seasonality typically indicate the presence or absence of statistically significant seasonality but do not provide a meaningful measure of its strength. Several statistical methods are available ranging from simple graphical techniques to more advanced statistical methods. Additionally, autocorrelation functions can be examined to assess regularity of periodicity or seasonality.



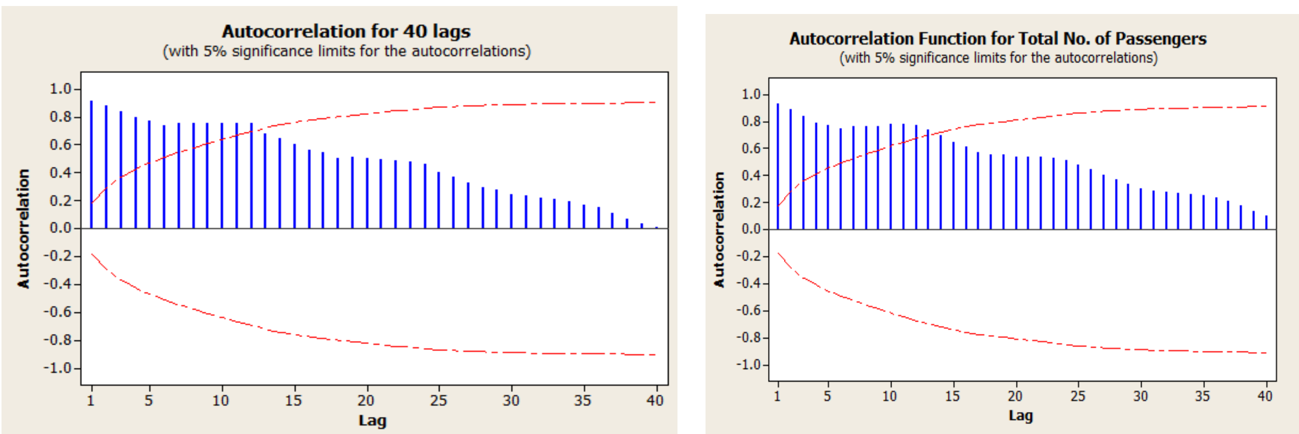




**Figure 5. Autocorrelation function for Gregorian and Hijri data**

**Seasonality for passenger time series data:** Seasonality is an important component of passenger numbers time series. The presence of predictable seasonality is a clue to the choice of an appropriate method for forecasting and planning purposes. Understanding seasonality is also essential for setting rational

## Scatter Diagrams

Scatter diagrams are drawn separately for the number of passengers for each year in Hijri and Gregorian dates as shown in Figures 3 and 4. Figure 3 seems to present approximately similar seasonality patterns for years 1428 to 1431, and for years 1435 and 1436. Figure 4 seems to present approximately similar seasonality pattern for years 2007 and 2008, for years 2009 and 2010, 2013 to 2015.

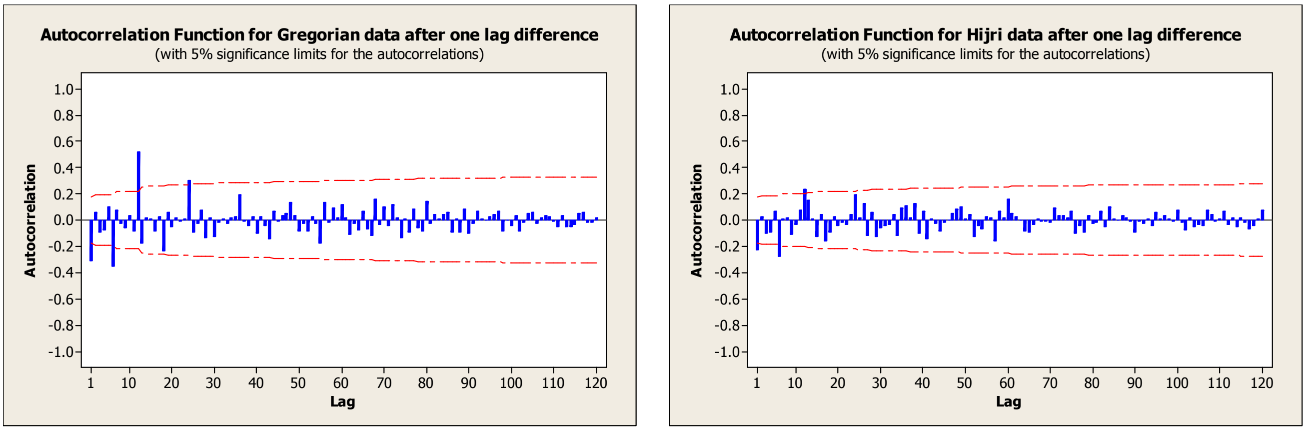



**Figure 6. Autocorrelation function for Gregorian and Hijri data after one lag differencing**

### Autocorrelation Functions

The correlation can refer to the relation between observation, e.g. between current observation and observation lagged by one or more time units. In this case all observations come from one variable, so a similarity between a given time series and a k-lagged version of itself over successive time intervals is called autocorrelation. The set of autocorrelation coefficients ($k$) arranged as a function of $k$ is the autocorrelation function (ACF). The graphical or numerical presentation of ACF is called autocorrelogram. The autocorrelogram presents not only autocorrelation coefficients but also the confidence intervals. If the autocorrelation coefficient is within the confidence interval, it is regarded as statistically insignificant. Minitab Software (http://www.minitab.com/en-us/products/minitab/free-trial/) is used to draw a graph of the autocorrelations for various lags, Figure 5 shows graphs of the autocorrelations for 40 lags for the Gregorian and Hijri data. Both figures are closely similar and show that there exists a significant relationship between successive time series values. The autocorrelation coefficients are large for the first several time lags, and then gradually drop toward zero as the number of lags increases, and the autocorrelation for time lag 1 is close to 1, for time lag 2 is large but smaller than for time lag 1 (Chandler and Scott, 2011). So, there is a trend pattern for both the data in Gregorian and Hijri calendars.

For seasonal pattern, a large correlation occurs at the first season lag and decreases over several seasonal lags. If there is seasonality, the ACF at the seasonal lag (12 for monthly data) will be large and positive. For seasonal monthly data, a large ACF value will be seen at lag 12 and possibly also at lags 24, 36, . . . The ACF peaks at lags 12, 24, 36, . . . , indicate seasonality of length 12. In addition, since both charts exhibit a similar pattern, we can fit the same model to both (Rob J Hyndman, 2018). The ACF shows an oscillation, indicative of a seasonal series, the peaks occur at lags of 12 months, because April 2011 correlates with April 2012, and 24 months, because April 2011 correlates with April 2013, and so on (Abbas Keshvani). The autocorrelations for 120 lags are plotted in Figure 6 for Gregorian and Hijri data using Minitab after one lag differencing to remove the trend. The ACF peaks at lags 12, 24, 36 indicate seasonality of length 12.

### The $x^2$ Goodness-of-Fit Test

Chi-Square goodness of fit test is used to find out how the observed value of a given phenomena is significantly different from the expected value.

In Chi-Square goodness of fit test, the term goodness of fit is used to compare the observed sample distribution with the expected probability distribution. In Chi-Square goodness of fit test, the null hypothesis assumes that there is no significant difference between the observed and the expected value so the data follow a specified distribution, whereas the alternative hypothesis assumes that there is a significant difference between the observed and the expected value so the data do not follow a specified distribution. The value of Chi-Square is calculated using the following formula:

$$\chi^2 = \Sigma \frac{(O-E)^2}{E}$$

Where $O$ denoted to observed value whereas $E$ denoted to expected value. The formula to calculate the expected value for a tabulated data of passengers is obtained as follows (Mario, 2014).

The probability that a person is in Row 1 and Column 1 =
= P(Row 1 and Column 1) = P(Row 1) * P(Column 1) = (Row Total/N) * (Column Total/N)

Expected Cell Frequency = The Probability * Total Number (N) =
= (Row Total/N * Column Total/N) . N = (Row Total * Column Total) / N

Table 3 represents the calculated expected value for the Gregorian data. Similar calculations are done for the Hijri data. The $x^2$ goodness-of-fit test is relatively popular for detecting seasonality because of its simple mathematical theory, which makes it easy to calculate and understand. This test requires a sample from a population with an unknown distribution function F(x) and a certain theoretical distribution function Fo(x). Although there is no restriction on the underlying distribution, usually the hypothetical distribution is a uniform distribution. The null hypothesis that there is no seasonal effect (i.e., Fo($x$) is a uniform distribution), then $E1 = E2 = \ldots = Ek$ (Eleazar, 2016). The decision rule depends on the level of significance and the degrees of freedom, df = (r-1)(c-1), where r and c are the numbers of rows and columns in the two-way data table. For r = 12, c = 11, then df = 110 and for level of significance = 95%, the critical value of Chi Square = 77.93.

**Table 3. The expected value for the Gregorian data**

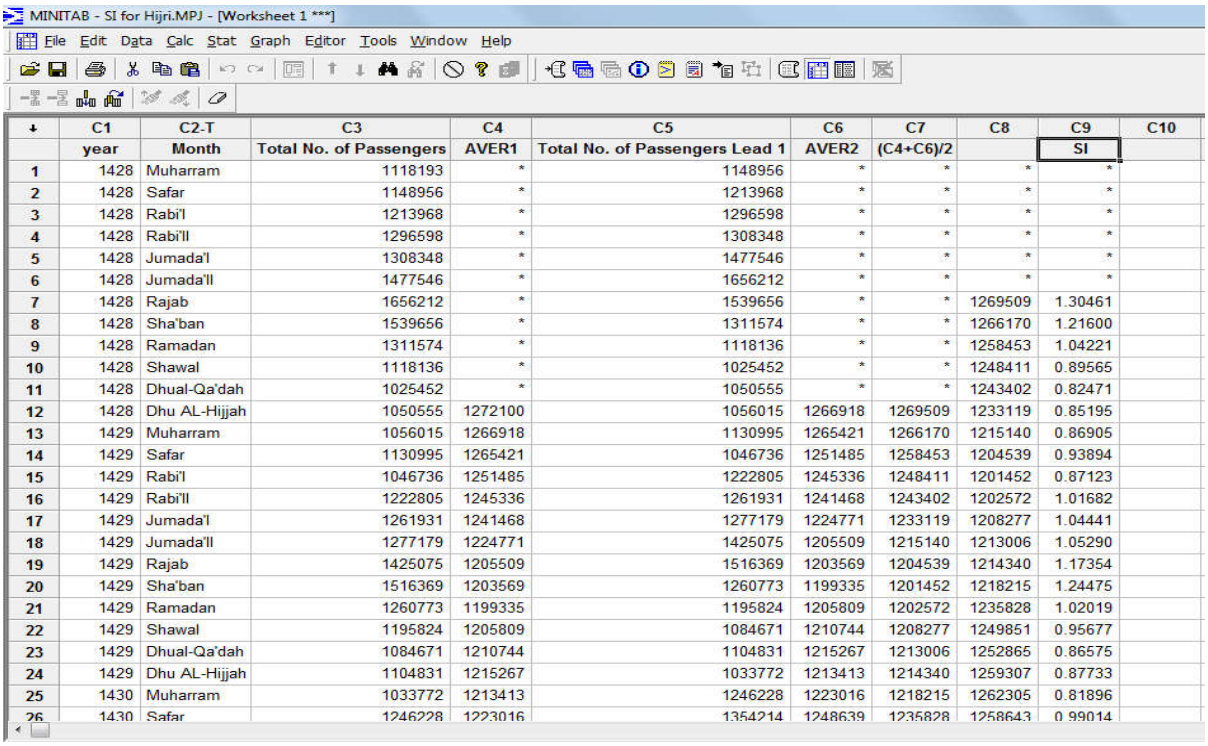
| 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1,293,300 | 1,250,364 | 1,318,797 | 1,389,375 | 1,583,086 | 1,832,345 | 1,940,768 | 2,128,489 | 2,269,672 | 2,351,107 | 1,432,637 |
| 1,233,100 | 1,192,163 | 1,257,410 | 1,324,703 | 1,509,398 | 1,747,054 | 1,850,431 | 2,029,413 | 2,164,024 | 2,241,669 | 1,365,951 |
| 1,401,295 | 1,354,774 | 1,428,921 | 1,505,393 | 1,715,280 | 1,985,352 | 2,102,829 | 2,306,225 | 2,459,197 | 2,547,432 | 1,552,267 |
| 1,402,622 | 1,356,057 | 1,430,274 | 1,506,819 | 1,716,904 | 1,987,233 | 2,104,821 | 2,308,410 | 2,461,526 | 2,549,845 | 1,553,737 |
| 1,418,681 | 1,371,583 | 1,446,649 | 1,524,070 | 1,736,561 | 2,009,985 | 2,128,919 | 2,334,839 | 2,489,708 | 2,579,038 | 1,571,526 |
| 1,417,079 | 1,370,034 | 1,445,016 | 1,522,349 | 1,734,600 | 2,007,715 | 2,126,515 | 2,332,202 | 2,486,897 | 2,576,126 | 1,569,751 |
| 1,492,452 | 1,442,905 | 1,521,875 | 1,603,321 | 1,826,862 | 2,114,503 | 2,239,622 | 2,456,249 | 2,619,172 | 2,713,147 | 1,653,245 |
| 1,319,560 | 1,275,752 | 1,345,574 | 1,417,586 | 1,615,230 | 1,869,550 | 1,980,175 | 2,171,707 | 2,315,756 | 2,398,845 | 1,461,726 |
| 1,202,363 | 1,162,446 | 1,226,067 | 1,291,682 | 1,471,773 | 1,703,505 | 1,804,305 | 1,978,826 | 2,110,081 | 2,185,790 | 1,331,902 |
| 1,148,958 | 1,110,815 | 1,171,609 | 1,234,311 | 1,406,403 | 1,627,842 | 1,724,165 | 1,890,934 | 2,016,360 | 2,088,706 | 1,272,744 |
| 1,103,218 | 1,066,592 | 1,124,967 | 1,185,172 | 1,350,413 | 1,563,037 | 1,655,525 | 1,815,655 | 1,936,087 | 2,005,553 | |
| 1,175,708 | 1,136,676 | 1,198,886 | 1,263,047 | 1,439,146 | 1,665,741 | 1,764,306 | 1,934,958 | 2,063,304 | 2,137,334 | |

**Table 4. Hijri Months' Orders**

| | Order of Month | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Year | Muhar | Safar | Rabi'l | Rabi'll | Juma'l | Juma'll | Rajab | Sha'b | Rama | Shawa | Dhul-Qa'dah | Dhul Hijjah |
| 1428 | 4 | 5 | 6 | 7 | 8 | 10 | 12 | 11 | 9 | 3 | 1 | 2 |
| 1429 | 2 | 5 | 1 | 7 | 9 | 10 | 11 | 12 | 8 | 6 | 3 | 4 |
| 1430 | 1 | 4 | 9 | 5 | 6 | 11 | 10 | 12 | 8 | 7 | 2 | 3 |
| 1431 | 1 | 5 | 6 | 4 | 7 | 3 | 10 | 11 | 12 | 9 | 2 | 8 |
| 1432 | 1 | 2 | 3 | 7 | 6 | 10 | 8 | 12 | 11 | 9 | 4 | 5 |
| 1433 | 2 | 5 | 7 | 8 | 6 | 9 | 10 | 12 | 11 | 3 | 1 | 4 |
| 1434 | 1 | 5 | 4 | 12 | 7 | 11 | 9 | 10 | 6 | 8 | 2 | 3 |
| 1435 | 1 | 2 | 9 | 6 | 11 | 7 | 8 | 12 | 4 | 10 | 5 | 3 |
| 1436 | 1 | 2 | 8 | 4 | 10 | 3 | 7 | 11 | 5 | 12 | 6 | 9 |
| 1437 | 1 | 2 | 10 | 3 | 4 | 12 | 8 | 9 | 6 | 7 | 11 | 5 |

**Table 5. Gregorian Months' Orders**

| | Order of Mongth | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Year | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
| 2007 | 2 | 3 | 5 | 7 | 8 | 2 | 11 | 12 | 9 | 6 | 1 | 4 |
| 2008 | 1 | 4 | 5 | 6 | 8 | 1 | 11 | 12 | 10 | 7 | 2 | 3 |
| 2009 | 1 | 3 | 5 | 10 | 8 | 1 | 12 | 11 | 9 | 6 | 2 | 4 |
| 2010 | 2 | 1 | 9 | 4 | 7 | 2 | 11 | 12 | 10 | 3 | 8 | 6 |
| 2011 | 2 | 1 | 4 | 8 | 7 | 2 | 11 | 11 | 10 | 5 | 3 | 6 |
| 2012 | 3 | 1 | 8 | 7 | 9 | 3 | 11 | 10 | 4 | 5 | 2 | 6 |
| 2013 | 6 | 3 | 12 | 7 | 10 | 6 | 4 | 8 | 1 | 5 | 2 | 9 |
| 2014 | 9 | 2 | 11 | 7 | 8 | 9 | 5 | 12 | 4 | 3 | 1 | 6 |
| 2015 | 10 | 1 | 11 | 4 | 7 | 10 | 9 | 12 | 5 | 3 | 2 | 8 |
| 2016 | 11 | 2 | 12 | 5 | 10 | 11 | 8 | 9 | 6 | 1 | 3 | 4 |




| ↓ | C1 | C2-T | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 |
|---|---|---|---|---|---|---|---|---|---|---|
| | year | Month | Total No. of Passengers | AVER1 | Total No. of Passengers Lead 1 | AVER2 | (C4+C6)/2 | | SI | |
| 1 | 1428 | Muharram | 1118193 | * | 1148956 | * | * | * | * | |
| 2 | 1428 | Safar | 1148956 | * | 1213968 | * | * | * | * | |
| 3 | 1428 | Rabi'l | 1213968 | * | 1296598 | * | * | * | * | |
| 4 | 1428 | Rabi'll | 1296598 | * | 1308348 | * | * | * | * | |
| 5 | 1428 | Jumada'l | 1308348 | * | 1477546 | * | * | * | * | |
| 6 | 1428 | Jumada'll | 1477546 | * | 1656212 | * | * | * | * | |
| 7 | 1428 | Rajab | 1656212 | * | 1539656 | * | * | 1269509 | 1.30461 | |
| 8 | 1428 | Sha'ban | 1539656 | * | 1311574 | * | * | 1266170 | 1.21600 | |
| 9 | 1428 | Ramadan | 1311574 | * | 1118136 | * | * | 1258453 | 1.04221 | |
| 10 | 1428 | Shawal | 1118136 | * | 1025452 | * | * | 1248411 | 0.89565 | |
| 11 | 1428 | Dhual-Qa'dah | 1025452 | * | 1050555 | * | * | 1243402 | 0.82471 | |
| 12 | 1428 | Dhu AL-Hijjah | 1050555 | 1272100 | 1056015 | 1266918 | 1269509 | 1233119 | 0.85195 | |
| 13 | 1429 | Muharram | 1056015 | 1266918 | 1130995 | 1265421 | 1266170 | 1215140 | 0.86905 | |
| 14 | 1429 | Safar | 1130995 | 1265421 | 1046736 | 1251485 | 1258453 | 1204539 | 0.93894 | |
| 15 | 1429 | Rabi'l | 1046736 | 1251485 | 1222805 | 1245336 | 1248411 | 1201452 | 0.87123 | |
| 16 | 1429 | Rabi'll | 1222805 | 1245336 | 1261931 | 1241468 | 1243402 | 1202572 | 1.01682 | |
| 17 | 1429 | Jumada'l | 1261931 | 1241468 | 1277179 | 1224771 | 1233119 | 1208277 | 1.04441 | |
| 18 | 1429 | Jumada'll | 1277179 | 1224771 | 1425075 | 1205509 | 1215140 | 1213006 | 1.05290 | |
| 19 | 1429 | Rajab | 1425075 | 1205509 | 1516369 | 1203569 | 1204539 | 1214340 | 1.17354 | |
| 20 | 1429 | Sha'ban | 1516369 | 1203569 | 1260773 | 1199335 | 1201452 | 1218215 | 1.24475 | |
| 21 | 1429 | Ramadan | 1260773 | 1199335 | 1195824 | 1205809 | 1202572 | 1235828 | 1.02019 | |
| 22 | 1429 | Shawal | 1195824 | 1205809 | 1084671 | 1210744 | 1208277 | 1249851 | 0.95677 | |
| 23 | 1429 | Dhual-Qa'dah | 1084671 | 1210744 | 1104831 | 1215267 | 1213006 | 1252865 | 0.86575 | |
| 24 | 1429 | Dhu AL-Hijjah | 1104831 | 1215267 | 1033772 | 1213413 | 1214340 | 1259307 | 0.87733 | |
| 25 | 1430 | Muharram | 1033772 | 1213413 | 1246228 | 1223016 | 1218215 | 1262305 | 0.81896 | |
| 26 | 1430 | Safar | 1246228 | 1223016 | 1354214 | 1248639 | 1235828 | 1258643 | 0.99014 | |

**Figure 7. Minitab screen for seasonality indexes for Hijri data**

The calculated Chi square for the Hijri and Gregorian data are 1,698,789 and 5,820,666 respectively which are greater than the critical value, then the data are seasonal for both.

**Month's Orders**

The months are arranged in an ascending ordered regarding to the no. of passengers. The month's order in every year will be indicated as shown in Tables 4 and 5 for Hijri and Gregorian months. Table 4 shows that seasonality is better in months: 1, 2 and 8 (Muharram, Safar and Sha'ban) than other months since their orders are the same for many years. Table 5 shows that seasonality is better in months: 2, 4, 5, 7, 8, 11 and 12 (February, April, May, July, August, November and December) since their orders are the same for 4 or more years. From both figures, it seems that Gregorian data is better in seasonality than that of Hijri data.

The seasonality indexes for Hijri data are calculated by Minitab using the method of ratio to moving average [20] as shown in Figure 7. The sum of 12 seasonality indexes for each year should be equal to 12. So, an adjustment process is done to calculate the adjusted seasonality indexes Figure 8 shows the adjusted seasonality indexes for Hijri data. From the figure, it can be concluded that there is a considerable variation for the values of each seasonality index. In the same manner, the seasonality indexes for Gregorian data are calculated and adjusted, results are shown in Figure 9. From the figures, it can be concluded that there is a considerable variation for the values of each seasonality index. Apparently, it seems that months1, 2, 5, 6, 9, 10 and 12 in hijri data and months 2, 3, 4, 5, 6, 9, 10, and 12 in Gregorian data are smoother than other months. So, Hijri data has 8 months with apparently smoother seasonality indexes, while Gregorian data has only 6 months.

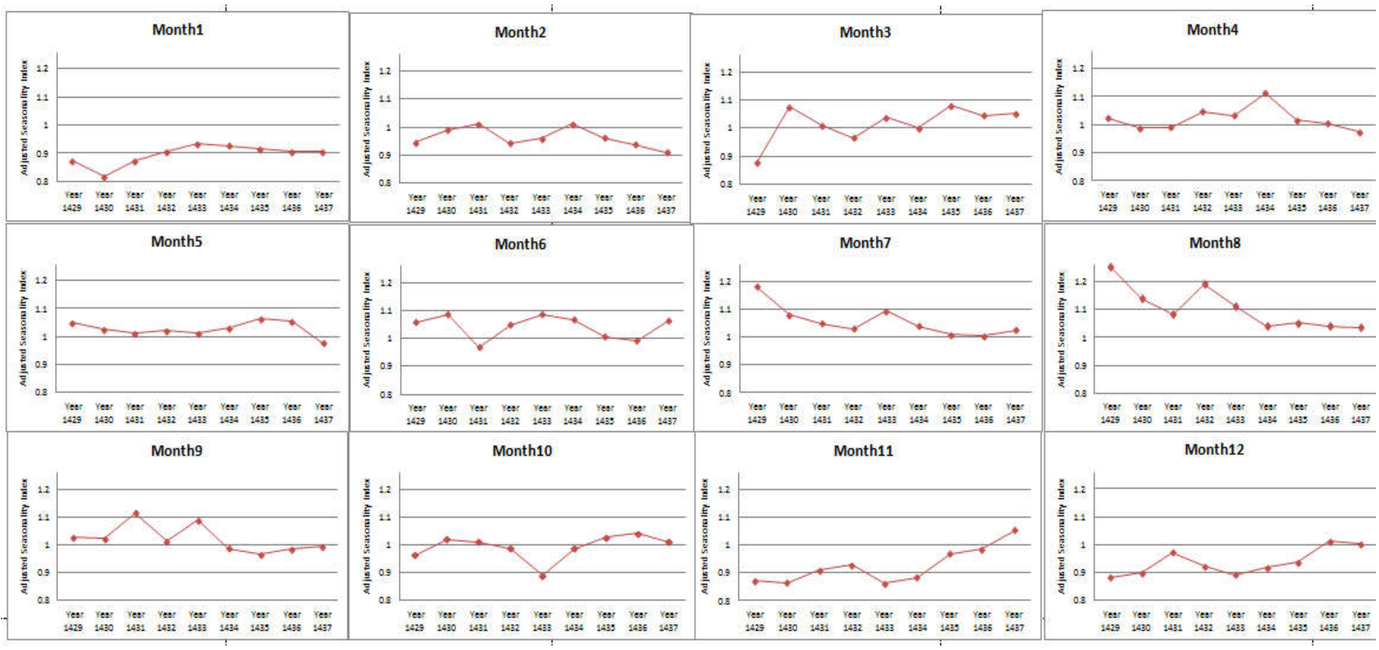



**Figure 8. Seasonality indexes graphs for the Hijri data**


**Figure 9. Seasonality indexes graphs for Gregorian data**

**Seasonality indexes**

But, it is hard to interpret data to others, and that's why we are in need to use statistics to assessing variations in the seasonality indexes.

The commonly used multiplicative model is of the form:

$$X_t = S_t * T_t * C_t * E_t$$

Where $X_t$ are the actual time series data, composed of seasonal ($S_t$), trend ($T_t$), cyclical ($C_t$) and error ($E_t$) components assumed for this study. Seasonal indexes ($S_t$) are used to indicate both the magnitude and direction of seasonal changes in the data. If there is no seasonality in the data, all indexes should equal 1, while if seasonality does exist the indexes vary around 1. It is assumed that if the mean seasonal indexes are equal, that they are all equal to one, thus indicating a nonseasonal series.

The question is to determine what amount of variation from 1 is significant. Analysis of Variance (ANOVA) is a statistical technique that can be used to test the hypothesis that the mean seasonal indexes for each period of a time series are equal, against that alternative that at least two of them differ (Kutner *et al.,* 2004).P-values were examined from the ANOVA tests to determine significance. A p-value less than 0.05 was considered to be evidence of a significant difference in the seasonal indexes, therefore, an indication of seasonality in the time series. The Minitab screen and the output are presented in Figures 10 and 11. A p-value less than 0.05 was considered to be evidence of a significant difference in the seasonal indexes for both the Gregorian and Hijri data, therefore, an indication of seasonality. The graph of the Confidence Interval (CI) for the seasonality indexes also clarify the existence of difference in their values.

**Results for: ANOVA-Gregorian.MTW**

**One-way ANOVA: C1, C2, C3, C4, C5, C6, C7, C8, C9, C10, C11, C12**

```
Source   DF       SS       MS      F      P
Factor   11  0.38134  0.03467  11.67  0.000
Error    96  0.28530  0.00297
Total   107  0.66664

S = 0.05452   R-Sq = 57.20%   R-Sq(adj) = 52.30%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level  N    Mean   StDev  -------+---------+---------+---------+--
C1     9  0.9782  0.1010              (----*----)
C2     9  0.9265  0.0314      (----*-----)
C3     9  1.0450  0.0535                       (----*----)
C4     9  1.0190  0.0283                   (-----*----)
C5     9  1.0376  0.0175                      (----*----)
C6     9  1.0337  0.0345                      (----*----)
C7     9  1.0910  0.0889                              (----*----)
C8     9  1.0804  0.0604                            (----*----)
C9     9  0.9893  0.0516               (----*----)
C10    9  0.9390  0.0355        (----*----)
C11    9  0.8995  0.0427  (----*-----)
C12    9  0.9461  0.0446         (----*----)
                          -------+---------+---------+---------+--
                              0.910     0.980     1.050     1.120
```

**Figure 10. ANOVA results for the seasonality indexes in Gregorian data**

```
Welcome to Minitab, press F1 for help.
```

**Results for: ANOVA-Hijri.MTW**

**One-way ANOVA: C1, C2, C3, C4, C5, C6, C7, C8, C9, C10, C11, C12**

```
Source   DF       SS       MS      F      P
Factor   11  0.35507  0.03228  12.63  0.000
Error    96  0.24538  0.00256
Total   107  0.60045

S = 0.05056   R-Sq = 59.13%   R-Sq(adj) = 54.45%

                          Individual 95% CIs For Mean Based on
                          Pooled StDev
Level  N    Mean   StDev  -------+---------+---------+---------+--
C1     9  0.8964  0.0360  (----*----)
C2     9  0.9623  0.0347            (---*----)
C3     9  1.0158  0.0638                   (----*----)
C4     9  1.0201  0.0407                    (----*----)
C5     9  1.0273  0.0263                     (----*----)
C6     9  1.0418  0.0421                       (----*----)
C7     9  1.0561  0.0554                         (----*----)
C8     9  1.1043  0.0766                                (----*----)
C9     9  1.0215  0.0491                    (----*----)
C10    9  0.9926  0.0452                (----*----)
C11    9  0.9245  0.0657      (----*----)
C12    9  0.9374  0.0477        (----*----)
                          -------+---------+---------+---------+--
                              0.910     0.980     1.050     1.120
```

**Figure 11. ANOVA results for the seasonality indexes in Gregorian data**

Conversely, if the mean seasonal indexes vary significantly from each other, and thereby from 1, a seasonal time series is indicated (Davey *et al.,* 1993).

**Plot changes in seasonal pattern:** To test for significance of the seasonal pattern, we calculate the observed changes in a time series based on the month of a year (the difference between number of passengers in a month and the previous month) (Bquanttrading, 2018). To help us visualize the results, we plot a time series for the changes in each month as shown in Figures 12 and 13. From the graphs, we can see that months: 2, 3, 7, 9 and 11are more smooth (with stronger seasonality) in Hijri data, and also months 2, 3, 7, 9 and 12 in Gregorian data. individual items of the distribution that can be explained by measures of dispersion. The most common measures of dispersion are Range, Quartile deviation, Mean deviation and standard deviation. Dispersion is the measure of the variation of the items. Measures of dispersion are meant to describe how spreads out data are.

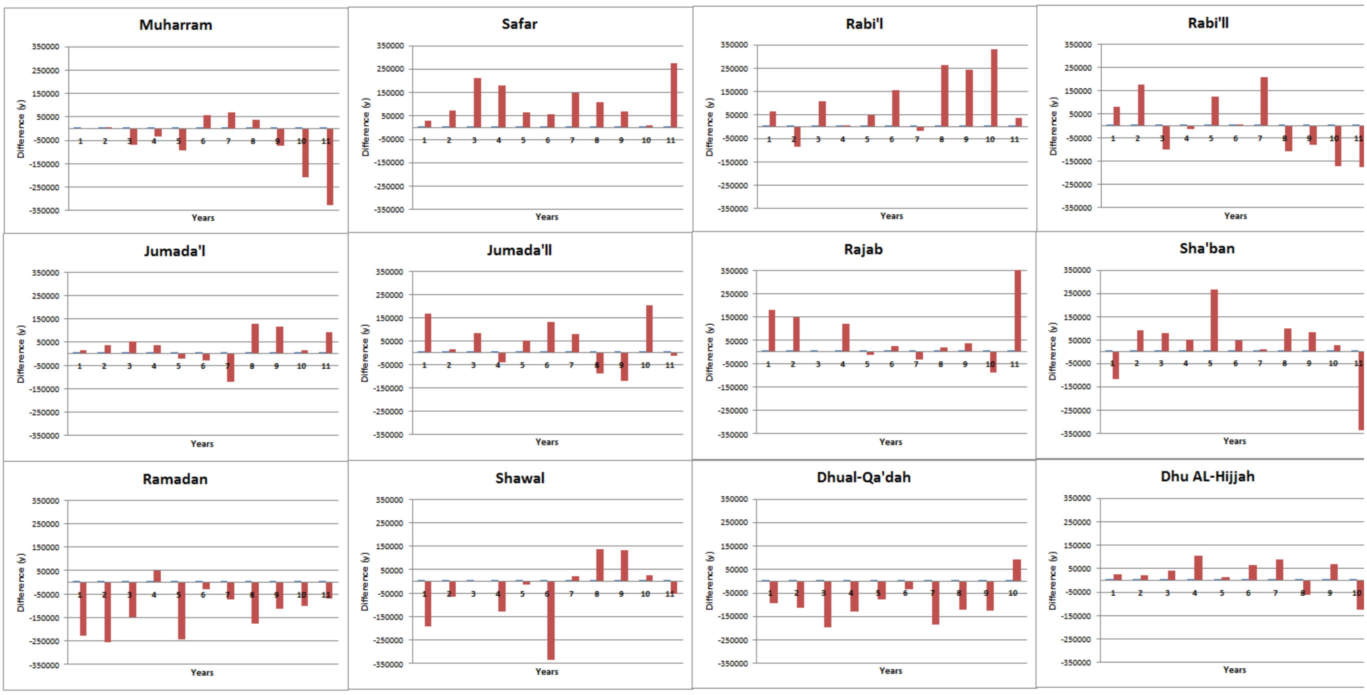



**Figure 12. Time series plot for the changes for the Hijri months**

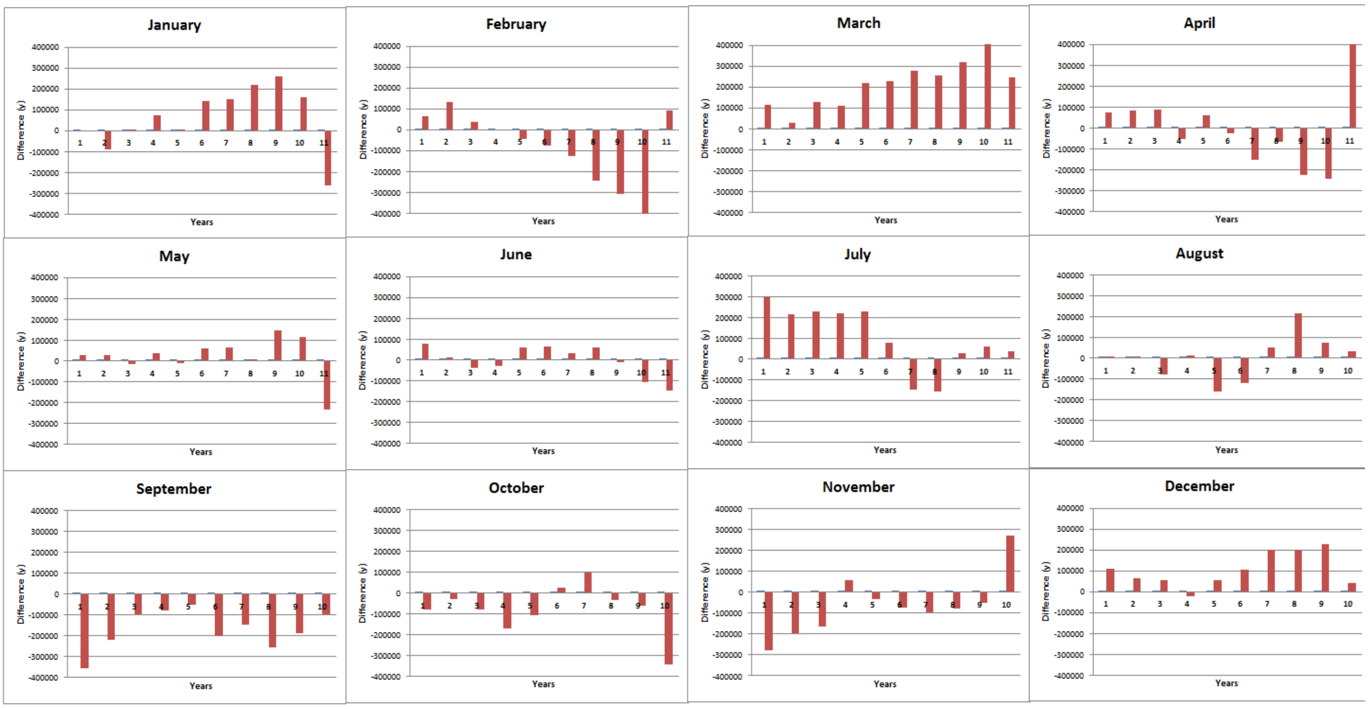



**Figure 13. Time series plot for the changes for the Gregorian months**

## Measures of dispersion

Collecting data can be sometimes easy, but it can be hard to interpret data to others. That’s why we use statistics. Two kinds of statistics are frequently used: measures of central tendency and dispersion. Measures of central tendency indicate the central position of a series. But they fail to reveal the degree of the spread out or the extent of the variability in

The more similar the scores are to each other, the lower the measure of dispersion will be. The less similar the scores are to each other, the higher the measure of dispersion will be. In general, and the more spread out a distribution is, the larger the measure of dispersion will be (Jayapriya, 2017). The range, inter-quartile range, mean deviation and standard deviation are all measures that indicate the amount of variability within a dataset. Small dispersion indicates high uniformity of the items, while large dispersion indicates less uniformity. The range and quartile deviation are not based on all observations. They are positional measures of dispersion. The range is the simplest measure of variability to calculate but can be misleading if the dataset contains extreme values. The inter-quartile range reduces this problem by considering the variability within the middle 50% of the dataset. The standard deviation is the most robust measure of variability since it takes into account a measure of how every value in the dataset varies from the mean (Durmuş Özdemir, 2016). Methods of studying dispersion are divided into two types**.** Mathematical methods where we can study the degree and extent of variation and graphic methods where we can study only the extent of variation, whether it is higher or lesser a Lorenz-curve is used. Mathematical methods commonly used as measures of dispersion are:

1. Range,
2. Quartile deviation,
3. Average deviation,
4. Standard deviation and coefficient of variation.

## 1. Ranges

Range is the simplest method of studying dispersion. In statistics, range represents the difference between the highest value of a data set and the lowest value of a data set. The range shows how spread out the values in a series are. If the range is a high number, then the values in the series are spread far apart; if the range is a small number, then the values in the series are close to each other. It is not a good measure of spread because it takes into account only the two extreme values, while computing range, we do not take into account frequencies of different groups (Sarah Boslaugh, 2012).

$$\text{Range} = x_{max} - x_{min} \qquad \text{(1)}$$

Where:
$x_{max}$ = The maximum value of the data set,
$x_{min}$ = The minimum value of the data set.

The ranges for the seasonal indexes are calculated for the seasonal indexes in Hijri and Gregorian calendars as shown in Table 6.

**Table 6. Ranges for seasonal indexes for Hijri and Gregorian data**

| Years | Months | | | | | | | | | | | | Total | Average |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | | |
| Hijri | 0.1159 | 0.10154 | 0.20417 | 0.21189 | 0.08584 | 0.11669 | 0.3020 | 0.215 | 0.14727 | 0.15071 | 0.22872 | 0.1588 | 2.03852 | 0.1699 |
| Gregorian | 0.2788 | 0.08438 | 0.15747 | 0.09784 | 0.05285 | 0.11409 | 0.3409 | 0.323 | 0.14501 | 0.13977 | 0.19232 | 0.1365 | 2.06259 | 0.1719 |

Table 6 shows that:

- Months: 1, 2, 5, 6, 9, 10 and 12 has smaller ranges than other months in Hijri data.
- Months: 2, 3, 4, 5, 6, 9, 10 and 12 has smaller ranges than other months in Gregorian data.
- Gregorian data has 9 seasonality indexes with less ranges than those of the Hijri data, while the Hijri data has 3 seasonality indexes with less ranges than those in Gregorian data.
- Hijri data has slightly smaller average ranges than Gregorian.

## 2. Quartile Deviations

Quartile- divides the distribution into four equal parts. Each set of scores has three quartiles. These values can be denoted by Q1, Q2, and Q3. First Quartile Q1 (lower quartile) is the middle number between the smallest number and the median of the data set (25th. Percentile). Second Quartile Q2 is the median of the data that separates the lower and upper quartile (50th. Percentile). Third Quartile Q3 (upper quartile) is the middle value between the median and the highest value of the data set (75th. Percentile). Quartile deviation is based on the lower quartile Q1 and the upper quartile Q3. The difference (Q3 - Q1) is called the inter quartile range. The difference Q3−Q1 divided by 2 is called semi-inter-quartile range or the quartile deviation (Gareth James, 2017). Thus

$$\text{Q.D} = (\text{Q3} - \text{Q1}) / 2 \qquad \text{(2)}$$

The quartile deviation is a slightly better measure of absolute dispersion than the range, but it ignores the observations on the tails.

There is an easier way to calculate the percentile directly in Microsoft Excel (2010) by using the built in function "PERCENTILE (array, k)" which returns the k-th. Percentile of values in a range. Q1 (P25) is calculated by using the Function "PERCENTILE (array, 0.25)" where the array is the seasonality indexes. Also, Q3 (P75) is calculated by using the function "PERCENTILE (array, 0.75)". The semi-inter-quartile range or the quartile deviation is the difference Q3−Q1 divided by 2. In the same manner, the quartile deviations for the Gregorian data is calculated. The results are shown in Table 7.

**Table 7. Quartile Deviations for Hijri and Gregorian data**

| Seasonality Index for Month | Quartile Deviation | |
|---|---|---|
| | Hijri | Gregorian |
| 1 | 0.0207 | 0.0828 |
| 2 | 0.0210 | 0.0200 |
| 3 | 0.0325 | 0.0284 |
| 4 | 0.0208 | 0.0170 |
| 5 | 0.0183 | 0.0119 |
| 6 | 0.0299 | 0.0211 |
| 7 | 0.0323 | 0.0799 |
| 8 | 0.0677 | 0.0347 |
| 9 | 0.0246 | 0.0424 |
| 10 | 0.0241 | 0.0213 |
| 11 | 0.0455 | 0.0212 |
| 12 | 0.0346 | 0.0354 |
| Total | 0.3720 | 0.4162 |
| Average | 0.0310 | 0.0347 |

Table 7 shows that:

- Months: 1, 2, 4, 5, 9 and 10 has smaller quartile deviation than other months in Hijri data.
- Months: 2, 3, 4, 5, 6, 10 and 11 has smaller standard deviation than other months in Gregorian data.
- Gregorian data has 9 seasonality indexes with less quartile deviation than those of the Hijri data, while the Hijri data has 3 seasonality indexes with less quartile deviation than those in Gregorian data.
- Hijri data has slightly smaller average quartile deviation than Gregorian.

## 3. Average Deviations

Average deviation is defined as a value which is obtained by taking the average of the deviations of various items from a measure of central tendency mean, median or mode, ignoring negative signs. Generally, the measure of central tendency from which the deviations arc taken, is specified in the problem. If nothing is mentioned regarding the measure of central tendency specified, then deviations are taken from median because the sum of the deviations (after ignoring negative signs) is minimum (Pham-Gia and T. L. Hung, 2011). The mean absolute deviation is more intuitively connected to the spread of the data set than what the standard deviation is. For day-to-day applications, the mean absolute deviation is a more tangible way to measure how spread out data are.

The Mean (Median) Absolute Deviation (MAD) =

$$MAD = \frac{1}{n}\sum |x_i - m| \qquad \text{……………………..} (3)$$

where: $x_i$ = The $i^{th.}$ Data number (observation),
n = Number of observations,
m = Mean, Median or mode.

The main steps to compute Average Deviation are:

i. Calculate the value of mean, median or mode, Take deviations from the given measure of central-tendency and they are shown as d.
ii. Ignore the negative signs of the deviation that can be shown as $|x_i - m|$ and add them to find $\sum|x_i - m|$.
iii. Apply formula (3) to get average deviation about mean, median or mode.

Instead of using the mentioned steps in getting the mean absolute deviation, we can get it directly in Microsoft Excel by using the built-in function "AVEDEV", it returns the average of the absolute deviations of data points from their mean. Table 8 shows the mean absolute deviation for the seasonality indexes in Hijri and Gregorian data.

**Table 8: Mean Absolute Deviation for Hijri and Gregorian data**

| Seasonality Index for Month | Mean Absolute Deviation (MAD) | |
|---|---|---|
| | Hijri | Gregorian |
| 1 | 0.0289 | 0.0817 |
| 2 | 0.0255 | 0.0240 |
| 3 | 0.0460 | 0.0401 |
| 4 | 0.0373 | 0.0238 |
| 5 | 0.0197 | 0.0142 |
| 6 | 0.0349 | 0.0301 |
| 7 | 0.0672 | 0.0901 |
| 8 | 0.0667 | 0.0704 |
| 9 | 0.0351 | 0.0458 |
| 10 | 0.0402 | 0.0290 |
| 11 | 0.0549 | 0.0391 |
| 12 | 0.0414 | 0.0390 |
| Total | 0.4978 | 0.5273 |
| Average | 0.0415 | 0.0439 |

Table 8 shows that:

- Months: 1, 2, 4, 5, 6 and 9 has smaller average deviation than other months in Hijri data.
- Months: 2, 4, 5, 6 and 10, 11 and 12 has average standard deviation than other months in Gregorian data.
- Gregorian data has 8 seasonality indexes with less average deviation than those of the Hijri data, while the Hijri data has 4 seasonality indexes with less average deviation than those in Gregorian data.
- Hijri data has slightly smaller total and average of the MAD than Gregorian data.

## 4. Standard Deviation

The standard deviation is extremely useful in judging the representativeness of the mean. The concept of standard deviation has a practical significance because it is free from all defects, which exists in a range, quartile deviation or average deviation. Standard deviation is calculated as the square root of average of squared deviations taken from actual mean. It is also called root mean square deviation. The square of standard deviation is called the variance. Using "STDEV" function in Microsoft Excel, the standard deviations for seasonal indexes were calculated for Hijri and Gregorian as shown in Table 9 (the colored cells show the lower values).

**Table 9. The Standard Deviations for seasonal indexes for Hijri and Gregorian data**

| Years | Months | | | | | | | | | | | | Total | Average |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | | |
| Hijri | 0.03566 | 0.03278 | 0.06127 | 0.05451 | 0.0263 | 0.04212 | 0.09436 | 0.08037 | 0.04678 | 0.0525 | 0.06949 | 0.05243 | 0.6486 | 0.0540 |
| Gregorian | 0.09803 | 0.03007 | 0.05003 | 0.03073 | 0.0172 | 0.03632 | 0.11231 | 0.09542 | 0.05267 | 0.0394 | 0.05544 | 0.04716 | 0.6648 | 0.0554 |

- Months: 1, 2, 5, 6, 9 and 12 has smaller standard deviation than other months in Hijri data.
- Months: 2, 4, 5, 6, 7, 10 and 12 has smaller standard deviation than other months in Gregorian.
- Gregorian data has 8 seasonality indexes with less standard deviations than those of the Hijri data, while the Hijri data has 4 seasonality indexes with less standard deviations than those in Gregorian data.
- Hijri data has slightly better average standard deviation than Gregorian data.

### Autoregressive to assess seasonality strength

A paper (Rahim Moineddin, 2003) propose the coefficient of determination of the autoregressive regression model fitted to the data ($R^2$) as a measure for quantifying the strength of the seasonality. The importance of this paper comes from the fact that the available statistical tests for seasonality typically indicate the presence or absence of statistically significant seasonality but do not provide a meaningful measure of its strength. For monthly data, one can use dummy variables for months in a regression model as a single predictor, and then, after correcting for the autocorrelation, calculate the coefficient of determination. When the time series is stationary and the trend is eliminated, the statistical significance of the dummy variables (months) indicates seasonality. The relationship between the stable seasonal factors and the estimates of the regression equation parameters are as follows: suppose there are $k$ years monthly, $n = 12\ k$, trend removed and centred (mean deleted) observations.

Let $\bar{y}$, $i$ = 1,2,...,12 denote the monthly average. The monthly averages, $\bar{y}s$, can be interpreted as crude estimates of stable seasonal factors, therefore, the range of parameter estimates is a good estimate of the magnitude of seasonal variation. For estimating one defines 11 dummy variables $m_i = 1$ if month equals $i$, $m_i = 0$ otherwise and then regress $m_i$ s on $y_t$ s. It is not difficult to show the ordinary least squared estimates of the parameters of the regression equation:

$$y_t = \beta_\circ + \beta_i m_i + \varepsilon_t,$$
$$\hat{\beta}_i = \bar{y}_i,\ i = 1, 2, \dots 11 \text{ and } \hat{\beta}o = \bar{y}_{12}$$

The estimated parameters $\hat{\beta}_i$ s can be used as seasonal factor estimates. Figure 14 presents the Excel output for the Gregorian data and similar output for the Hijri data is obtained. $R^2$ = 0.38 and 0.29 for the Gregorian and Hijri data respectively, which means moderate and week seasonality. The p-values indicate that coefficients for months 9 and 11 in Hijri data and for months 2, 3, 10 and 11 in Gregorian data.

| SUMMARY OUTPUT | | | | | | |
|---|---|---|---|---|---|---|
| *Regression Statistics* | | | | | | |
| Multiple R | 0.61764868 | | | | | |
| R Square | 0.38148989 | | | | | |
| Adjusted R Square | 0.3218091 | | | | | |
| Standard Error | 133753.818 | | | | | |
| Observations | 126 | | | | | |
| | | | | | | |
| ANOVA | | | | | | |
| | *df* | *SS* | *MS* | *F* | *Significance F* | |
| Regression | 11 | 1.25792E+12 | 1.14E+11 | 6.3921713 | 3.58171E-08 | |
| Residual | 114 | 2.03947E+12 | 1.79E+10 | | | |
| Total | 125 | 3.29739E+12 | | | | |
| | | | | | | |
| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* |
| Intercept | 93367.8 | 42296.67121 | 2.20745 | 0.0292853 | 9578.420103 | 177157.1799 |
| X Variable 1 | -38899.6 | 59816.52607 | -0.65032 | 0.516797 | -157395.6774 | 79596.47743 |
| X Variable 2 | -184829.891 | 58441.24968 | -3.16266 | 0.0020038 | -300601.5564 | -69058.22541 |
| X Variable 3 | 116830.927 | 58441.24968 | 1.999118 | 0.0479746 | 1059.261775 | 232602.5928 |
| X Variable 4 | -103565.709 | 58441.24968 | -1.77213 | 0.0790438 | -219337.3746 | 12205.95641 |
| X Variable 5 | -84108.8 | 58441.24968 | -1.4392 | 0.1528333 | -199880.4655 | 31662.8655 |
| X Variable 6 | -107434.891 | 58441.24968 | -1.83834 | 0.0686156 | -223206.5564 | 8336.774588 |
| X Variable 7 | -5766.8 | 58441.24968 | -0.09868 | 0.9215681 | -121538.4655 | 110004.8655 |
| X Variable 8 | -101501.9 | 59816.52607 | -1.69689 | 0.0924467 | -219997.9774 | 16994.17743 |
| X Variable 9 | -275590.8 | 59816.52607 | -4.60727 | 1.07E-05 | -394086.8774 | -157094.7226 |
| X Variable 10 | -182908.1 | 59816.52607 | -3.05782 | 0.0027785 | -301404.1774 | -64412.02257 |
| X Variable 11 | -171774.1 | 59816.52607 | -2.87168 | 0.0048694 | -290270.1774 | -53278.02257 |

**Figure 14. Excel output for the Gregorian data**

**Table 10. Seasonality results for Gregorian and Hijri data**

| # | Method | | Gregorian | Hijri |
|---|---|---|---|---|
| 1 | Scatter plots | | √ | √ |
| 2 | Autocorrelation Functions | | √ | √ |
| 3 | $x^2$ Goodness-of-Fit Test | | √ | √ |
| 4 | Month's Orders | | √ | |
| 5 | Seasonality Indexes using Ratio to Moving Average | | √ | √ |
| 6 | Plot of Changes in the Seasonal Pattern | | √ | √ |
| 7 | Measures of Dispersion for Seasonality Indexes | Ranges for Seasonality Indexes | √ | √ |
| 8 | | Quartile Deviations for Seasonality Indexes | √ | |
| 9 | | Average Deviation for Seasonality Indexes | √ | |
| 10 | | Standard Deviation | √ | |
| 11 | Autoregression to Assessing the Strength of Seasonality | | √ | |

**Table 11. Summary of results for Hijri data**

| # | Method | | Seasonality for Month: 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Month's Orders | | √ | √ | | | | | | √ | | | | |
| 2 | Seasonality Indexes using Ratio to Moving Average | | √ | √ | | | √ | √ | | | √ | √ | | √ |
| 3 | Plot of Changes in the Seasonal Pattern | | | √ | √ | | | | √ | | √ | | √ | |
| 4 | Measures of Dispersion for Seasonality Indexes | Range | √ | √ | | | √ | √ | | | √ | √ | | √ |
| 5 | | Quartile Deviation | √ | √ | | √ | √ | | | | √ | √ | | |
| 6 | | Average Deviation | √ | √ | | | √ | √ | | | √ | | | |
| 7 | | Standard Deviation | √ | √ | | | √ | √ | | | √ | | | √ |
| 8 | Autoregression to Assess Strength of Seasonality | | | | | | | | | | √ | | √ | |

**Table 12. Summary of results for Gregorian data**

| # | Method | | Seasonality for Month: 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Month's Orders | | | √ | | √ | √ | | √ | √ | | | √ | √ |
| 2 | Seasonality Indexes using Ratio to Moving Average | | √ | √ | √ | √ | √ | | | √ | √ | | √ | √ |
| 3 | Plot of Changes in the Seasonal Pattern | | | √ | √ | | | | √ | | √ | | | √ |
| 4 | Measures of Dispersion for Seasonality Indexes | Range | | √ | √ | √ | √ | √ | | | √ | √ | | √ |
| 5 | | Quartile Deviation | | √ | √ | √ | √ | √ | | | | √ | √ | |
| 6 | | Average Deviation | | √ | | √ | √ | √ | | | | √ | √ | √ |
| 7 | | Standard Deviation | | √ | | √ | √ | √ | √ | | | √ | | √ |
| 8 | Autoregression to Assess Strength of Seasonality | | | √ | √ | | | | | | | √ | √ | |

## RESULTS

Table 10, 11 and 12 summarizes the results obtained from all the applied methods for Hijri and Gregorian data respectively. From Table 10 it is clear that Gregorian data is better in seasonality than Hijri data. For the Hijri data (Table 11), we can see that months 1, 2, 5, and 9 have better seasonality and months 3, 4, 7, 8 and 11 are the worst. For the Gregorian data (Table 12), we can see that months 2:6 and 10:12 have better seasonality and months 1 and 8 are the worst.

### Conclusion

Monthly data for the number of passengers in Saudi Arabian Airlines are usually collected in Gregorian and partially in Hijri calendars. Since some data is not available, an approximate method based on the average number of passengers per day is used to transform the data from Gregorian to Hijri calendar. For studying seasonality, several statistical methods are available ranging from simple graphical techniques to more advanced statistical methods. These methods failed to answer the question of comparing seasonality between two sets of time series and to determine which data is better than the other in expressing seasonality pattern. Another important open question is to identify months of the year having better seasonality than others. To explore the seasonality patterns for the time series data in Gregorian and Hijri calendars we used the following methods: Scatter Diagrams, Autocorrelation Functions, $x^2$ Goodness-of-Fit Test, Seasonality Indexes using Ratio to Moving Average, Plot of Changes in the Seasonal Pattern and Autoregression to assess strength of seasonality. New methods are proposed to assess the seasonality of the data and moreover to identify which months of the year have better seasonality, these methods are: Month's Orders, Measures of Dispersion for Seasonality Indexes (Range, Quartile Deviation, Average Deviation and Standard Deviation). The results obtained from all the applied methods clarify that Gregorian data is better in seasonality than Hijri data. For the Hijri data, months 1, 2, 5, and 9 have better seasonality and months 3, 4, 7, 8 and 11 are the worst. For the Gregorian data, months 2:6 and 10:12 have better seasonality and months 1 and 8 are the worst.

## REFERENCES


Abbas Keshvani, How to use autocorrelation function (ACF) to determine seasonality, *Abbas Keshvani website under Time SeriesAbbas Keshvani under Time Series* https://coolstatsblog.com/2013/08/11/how-to-use-autocorreation-function-acf-to-determine-seasonality/

Bquanttrading, 2018. Testing for seasonality in Excel, Posted on June 29, 2015. Acced at 8 February. https://quantmacro.wordpress.com/2015/06/29/testing-for-seasonality-in-excel/

Chandler and Scott, Statistical methods for trend Detection and analysis in the environmental sciences, Wiley, 2011.

Davey, A.M. and B. E. Flores, B.E., Identification of seasonality in time series: A note, Math. Comput. Modelling, Vol. 18, No. 6, pp. 73-81, 1993.

David R. Anderson, Dennis J. Sweeney, Thomas A. Williams, Jeffrey D. Camm and James J. Cochran, Statistics for Business & Economics, 12th. Edition, South Western Cengage Learning, 2008.

Department of Network Planning and Scheduling, Network & Revenue Management Division, Commercial Function, Saudi Arabian Airlines.

Durmuş Özdemir, Applied Statistics for Economics and Business, Second edition, ISBN-10 3319264958, Springer, 2016.

Eleazar C. Nwogu, Iheanyi S. Iwueze and Valentine U. Nlebedim, 2016. Some Tests for Seasonality in Time Series Data, *Journal of Modern Applied Statistical Methods.*

Eleazar C. Nwogu, Iheanyi S. Iwueze and Valentine U. Nlebedim: Some Tests for Seasonality in Time Series Data, Journal of Modern Applied Statistical Methods 2016.

Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, An Introduction to Statistical Learning: with Applications in R, Springer Texts in Statistics, 7th. Ed., 2017.

International Astronomical Center webpage, visited at 3 January, 2018: http://www.icoproject.org/conv#togre

Jain T.R., Ashok Gupta and Sat Pal Gupta, Quantitative methods, Educational Publishers, 2007.

Jayapriya, J. Measures of dispersion and its interpretation, Research maGma, *An International Multidisciplinary Journal,* Volume-I, Issue-III, May-2017, ISSN NO- 2456-7078, 2017. http://researchmagma.com/admin/articles/Jayapriya.pdf

John E. Hanke and Dean Wishern, Business forecasting, Ninth Edition, Pearson, 2014.

Jones RH, Ford PM, Hamman RF. Seasonality comparisons among groups using incidence data. *Biometrics* 1988, 44:1131-1144.

Kutner, M., Nachtsheim, C., Neter, J., and Li, W. (2004). Applied Linear Statistical Models, McGraw-Hill/Irwin, Homewood, IL

Lorenz curve graphing tool & Gini coefficient calculator website, visited at January, 1, 2018. http://www.peterrosenmai.com/lorenz-curve-graphing-tool-and-gini-coefficient-calculator.

Mario F. Triola, Essentials of Statistics, 5th. Edition, Pearson, 2014.

Marrero O. 1983. The performance of several statistical tests for seasonality in monthly data. *J Statist Comput Simulation,* 17:275-296.

Microsoft Office, Microsoft Excel (Version 2010), Microsoft webpage, visited on January, 5, 2018. [Computer software]. https://microsoft-excel.en.softonic.com/.

Minitab software website, visited at 23 December 2017: http://www.minitab.com/en-us/products/minitab/free-trial/

Pham-Gia T. and T. L. Hung, The mean and median absolute deviations, Mathematical and Computer Modelling, 34 (2001) pp. 921-936, Elsevier Science Ltd. https://doi.org/10.1016/S0895-7177(01)00109-1.

Rahim Moineddin, Ross EG Upshur, Eric Crighton and Muhammad Mamdani, Autoregression as a means of assessing the strength of seasonality in a time series, Population Health Metrics, 2003. Available from: https://doi.org/10.1186/1478-7954-1-10

Rao Vallabhaneni S., Wiley CIAexcel Exam Review, Part 3: Internal Audit Knowledge Elements, John Wiley & Sons Inc., 2017, ISBN: 978-1-119-43921-9.

Rob J Hyndman and George Athanasopoulos, Forecasting Principles and Practice, 2nd edition, Springer, 2016. http://otexts.org/fpp2/

Rob J Hyndman, Forecasting using R, Autocorrelation and seasonality, Revolution Analytics, Monash University. Accessed at 2 January 2018: https://robjhyndman.com/talks/RevolutionR/3-Seasonality.pdf

Robert H. Shumway and David S. Stoffer, Time Series Analysis and Its Applications: With R Examples, 4th. ed. (Springer Texts in Statistics), 2017.

Sarah Boslaugh, Statistics in a Nutshell, 2nd Edition, O'Reilly Media, Inc. ISBN: 9781449361129, 2012.

UBC Real Estate Division, 2011. Understanding Data Patterns and Forecasting Techniques, Access date, December 27, 2017, https://professional.sauder.ubc.ca/re_creditprogram/course_resources/courses/content/460/materials/R2B46011_lesson07.pdf