**NU-590**

**ASSIGNMENT 2: DESCRIPTIVE STATISTICS**

**EXCEL EXERCISE [10 Points]**

We will be working on the “Dataset NHAMCS” spreadsheet from Canvas in order to complete the following exercises.

At the end of this session, you will learn the following:

* Sort and filter data
* Creating a frequency distribution table using the COUNTIF function
* Using the IF function to create new variables
* Use of the F4 function
* Using the Data Analysis Tool Pak to provide descriptive statistics
  + Mean, Median, Mode
  + Range
  + Standard Deviation

**Part 1. [10 Points]**

1. Sort the Data by Reason for Visit [RFV1]
2. Use the Filter and COUNTIF function in EXCEL and report the number of ER visits due to the following reasons:
   1. Abdominal Pain, cramps, spasms, NOS
   2. Accident
   3. Back pain, ache, soreness, discomfort
   4. Chestpain
   5. Earache
   6. Fever
3. Create a frequency distribution table using the **COUNTIF** function in excel:

|  |  |
| --- | --- |
|  | Number of visits |
| Abdominal Pain, cramps, spasms, NOS |  |
| Accident |  |
| Back pain, ache, soreness, discomfort |  |
| Chestpain |  |
| Earache |  |
| Fever |  |

1. Using the IMMEDR variable, report how many of the abdominal pain and chest pain cases were “Emergent”, “Urgent”, “Semi-urgent”.

Part 2.

Using the Data Analysis Tool Pak to provide descriptive statistics

* + Mean, Median, Mode
  + Range [Minimum, Maximum]
  + Variance
  + Standard Deviation
  + Skewness

1. Go to File, Options, Add-ins, Analysis Tookpak –VBA

Under Manage, select Go. Select Analysis Tookpak.

When you select “Data” tab, at the extreme right you will see “Data Analysis” appear.

1. Using this, provide the mean age of those patients visiting the ED in the United States in 2016.
   1. You will need to transform some of the data in the Age variable. Using “Find and Replace” replace all those observations with “Under one year” as 0 and those “93 years and over” as 93.
   2. If the data doesn’t read “Age” as a numeric variable, make sure that you transform the variable into a Numeric variable. [Hint: Use the Value function we learnt under Assignment 1].
   3. Using the Data Analysis Toolpak, provide the descriptive summary of the “Age Variable” including mean, median, range, variance etc.. Describe the age distribution in 3-4 sentences.
2. Provide the frequency distribution of the following age categories. Calculate the percentage and cumulative percentages associated with each of the categories.

|  |  |  |  |
| --- | --- | --- | --- |
| **Age group** | **Frequency** | **Percentage** | **Cumulative percentage** |
| Under 15 years |  |  |  |
| 15-24 years |  |  |  |
| 25-44 years |  |  |  |
| 45-64 years |  |  |  |
| 65-74 years |  |  |  |
| 75 years and over |  |  |  |

* Save the dataset on your thumb drive.