simplilearn
your pace, your place

Analyze the Healthcare Cost and Utilization in Wisconsin Hospitals

Business Analytic Foundation with SAS Tools & Excel- Question

Question

A nationwide survey of hospital costs conducted by the US Agency for healthcare consists of hospital records of inpatient samples. The given data is restricted to the city of Wisconsin and is related to patients in the age group 0-17 years. The Agency wants to analyze the data to research on the healthcare costs and their utilization.

Here is a detailed description of the given dataset:

AGE	: Age of the patient discharged
FEMALE	: Binary variable that indicates if the patient is female
LOS	: Length of stay, in days
RACE	: Race of the patient (specified numerically)
TOTCHG	: Hospital discharge costs
APRDRG	: All Patient Refined Diagnosis Related Groups

To complete this project, you will require a strong understanding of the following concepts:

- Lesson 3 – Histogram, Summaries, ANOVA
- Lesson 4 – Linear Regression

The goals of this project are:

- To record the patient statistics, the agency wants to find the age category that frequents the hospital and has the maximum expenditure.
- In order of severity of the diagnosis and treatments and to find out the expensive treatments, the agency wants to find the diagnosis related group that has maximum hospitalization and expenditure.
- To make sure that there is no malpractice, they need to analyze if the race of the patient is related to the hospitalization costs.
- To properly utilize the costs, the agency has to analyze the severity of hospital costs by age and gender for proper allocation of resources.
- Since the length of stay is the crucial factor for inpatients, the agency wants to find if the length of stay can be predicted from age, gender, and race.
- To perform a complete analysis, the agency wants to find the variable that mainly affects the hospital costs.

The data can be downloaded from the URL mentioned below (under the name HospitalCosts):

http://instruction.bus.wisc.edu/jfrees/jfreesbooks/Regression%20Modeling/BookWebDec2010/data.html

The total time provided to complete this task is 2 hours.