Stat Project

You are an employee of a major technology company that produces hardware and software.  The company is in a process of studying the salary structure of sales representatives in order to compete with other firms in the industry. As a data analyst, you are asked by human resource manager to analyze the factors that determine the salary of sales representatives in hardware and if there is a significant difference in salary between male and female employees.

With the help of IT group, a data set of over 20,000 records was extracted from the enterprise data warehouse.  The description of the data set is given below:

Variable​Description

Employee:​Unique ID for each sales representative

Business:​One of the two product groups: Hardware and Software

Age:​           Employee’s actual age

GEN:          Female​1 for female, 0 otherwise

Years:​The number of years the employee has been employed at the company

College​:   Whether or not the employee has a 4-year college degree (Yes/No)

Personality: ​Four personality types based on the Myers-Briggs Personality assessment:

Analyst: This personality type exemplifies rationality. Analysts tend to be open-minded and strong-willed. They like to work independently and usually approach things from a very practical perspective.

Diplomat: Diplomats care about people and tend to have a lot of empathy toward others. They exemplify cooperation and diplomacy.

Explorer: Explorers are highly practical and can think on their feet. They tend to be very good at making quick, rational decisions in difficult situations.

Sentinel: Sentinels are cooperative and practical. They like stability, order, and security. People with this personality type tend to be hard working and meticulous.

Certificates​The number of relevant professional certifications each employee has earned

Feedback​The average feedback score that each employee receives from his or her peers and supervisor on the 360-degree annual evaluation. The possible scores range from 0 (lowest) to 4 (highest).

Salary:​Annual base salary of each employee

NPS​The net promoter score (NPS) is a key indicator of customer satisfaction and loyalty.

To answer ALL of the questions in this item set, the data set must be filtered to contain observations that have the following THREE characteristics:

.1 The sales representative must work in the hardware industry.

2. The sales representative must have a college degree.

3. The sales representative must be 40 years, or older

Therefore, the first task is to filter the TechSales\_Reps data set to get the appropriate subset of observations.

You should verify that there are 21,988 observations in the filtered data set.

There are 21,988 observations in the filtered data that includes the Hardware industry, reps with a college degree, and reps who are 40 years or older than 40 years. Copy these observations to a separate worksheet/data frame in order to answer the following questions.

I. Descriptive statistics

Use Summary measures to examine the salaries and net promoter scores (NPS) of sales representatives depending on their personality types and sex in the hardware group.

.

II. Contingency Table.

a. Construct a contingency table that cross-classifies the data by gender and Personality. Provide the frequencies in the accompanying table.

b. Graph: Row = Gender and Column Personality

c. How many sales representatives are female and have a Diplomat personality type?

c. How many sales representatives are non-female and have a Sentinel personality type?

III. Two Sample t-test

a) Determine if there is a significant difference in:

i. Salary

ii. Age

iii. Feedback

iv. Number of certificate

Between female and non-female.

IV. Regression

Estimate the following model:

Salary = β0 + β1Certificates + β2NPS + β3Female + β4Analyst + β5Diplomat + β6Explorer + ε

Dummy variables

(Female) If Female = 1

0 other wise

(Analyst)If Analyst = 1

0 other wise

(Diplomat) If Diplomat = 1

       = 0 otherwise

(Explorer)If Explorer = 1

                = 0 otherwise

a. Test the significance of the variables

b. Interpret  the F- Static

c. Interpret the R-square

d. How would you interpret β3?

e. What factors appear to influence the salary?

Write a report of your findings that you present to the Humana Resource Manager.  The report should be at least five pages.