

**ANL252**

**Python for Data Analytics**

**Tutor Marked Assignment (TMA)**

**Submitted by:**

|  |  |
| --- | --- |
| **Name** | **PI No.** |
| Alexandre Oudea | K1882534 |

**Date Submitted: 13/08/2022**

Question 1a)

As stipulated by the question the 2 charts and respective tables in the snapshots below are based on correlation analysis which was done in part 1b).

Generally, an organization focuses on the performance of its staff as it serves as a major indicator of productivity. The focus of this datasets study is thus centered around the variable “PerformanceScore” which is assumed to be a form of KPI (key performance Index) that measures if workers meet their target objectives in the workplace.

What was discovered in the course of this exercise was that, relatively, only 2 variables stand out compared to the others that share a somewhat distinct correlation with PerformanceScore (the target variable). They are; Satisfaction (0.2) and Survey (0.4). Survey\_int is a column variable that rounds the Survey variable to the nearest whole number for presentation purposes on the clustered bar chart. However, during the analysis Survey was used instead due to the higher precision of its values.

The charts on excel were created using the pivot table and chart function to acquire a cross tab frequency distribution showing each performance category to each satisfaction and survey score

Below we can observe from the 2 charts that generally, and somewhat as expected, the majority of staff tend to meet their objectives at work. Most of the staff have moderate or higher than moderate sentiments in terms of contentment as well as engagement. Based purely on these charts, one could conclude that because the staff is engaged and satisfied at work they perform better.

On the other hand, what can be further gathered from correlation analysis in Jupyter Notebook is the nature of the relationship between satisfaction and engagement at work and performance. Surprisingly, the correlation is not particularly strong unlike at first glance in the frequency distribution charts. While the polarity of the coefficient is positive which is expected, the strength of correlation is generally weak. From here we can thus conclude that engagement and happiness at work generally share a direct relationship with a staff workers performance compared to the rest of the variables which don’t seem to have an impact (based only on this data set).

Moreover, we can also note that engagement scores are more determinant than satisfaction scores. Considering the two surveys can be assumed to be somewhat similar, it is interesting to note that the two do not share a strong correlation to each other either. Hence, we can not only conclude that engagement matters more than happiness towards better performance but also that just because a staff is engaged does not imply that they are happy. Just as a happy staff member themselves may not be engaged.

Note that none of the newly derived variables created in Jupyter Notebook have any clear correlation to the PerformanceScores of the staff which in itself is useful to note.

Moreover, further analysis would show that beyond Performance levels, other potential target variables do not share strong relationships to typical factors. For instance, satisfaction and salary variables do not share a strong correlation when generally speaking salary should affect satisfaction. What is likely as a possible explanation is the attitude with which the survey’s for both satisfaction and engagement were biased. As an evaluating point, it is possible that most of the staff are reluctant to make an honest review of their positions and environment for fear of repercussion. It may also be possible that the attitude with which the survey was completed may have been lack-luster or lazy on the part of the staff. Reasons could include that the organization has a poor reputation of addressing concerns.

