# Analysis and Insights of Employee Data

## Problem Statement

Data is given in a Excel file which contains a denormalized table. The data is required to be normalized and modelled in a star schema in Power BI as a tool. Develop and interactive dashboard that is aimed at delivering insights and answer the questions that has been put forward.

## Methodology

Before bringing in the data into Power BI a quick verification and inspection of the data must be done preferably as close to source as possible in this case the source being an Excel file. The data was inspected for missing values and inconsistencies. Data verification and validation is an important step in producing accurate insights. Once this was done the data was imported into Power BI and correct datatypes was assigned to each column if it was not correct. Next the denormalized table given was normalized by splitting into various dimension tables and one fact table. The dashboard was then constructed, and analysis was done based on the data. It was assumed that the performance score is used a metric for productivity, and this was correlated using th correlation measure in Power BI.

# *Insights and Answers to Questions*

**1.Departmental Performance: Which department has the highest average performance score, and how does it compare to other departments?**

The department with the best performance score was Engineering. The Department performance scores ranges from 2.98 to 3.02. These scores would indicate that on average the departments are performing at almost at the same level. Comparing the Engineering department to the other departments we see that Operations department is very close with a performance score of 3.01.

**2.Job Title Analysis: How does the average salary vary across different job titles, and which job title has the highest average performance score?**

There is a range of salaries which start from $4500 up to $7800. The highest average salary by Job Title is an Engineer at $7800 and the lowest salary is a technician at $4500. The average company salary is $6400 and there are 3 jobs which fall below the salary average which Specialist, Analyst and Technician. The highest performance score is 3.01which is the specialist position. It is observed that the performance scores for the different positions ranges from 2.98 to 3.01 which indicates that on average the different positions are performing at the same level.

**3.Education Impact: Does education level impact employee satisfaction, and if so, which education level has the highest satisfaction score?**

It can be observed that the Employee satisfaction ratings range from 2.99 to 3.00. This indicates that across the education levels there are no significant differences in satisfaction levels. The highest satisfaction score is for the High School level of education.

**4.Productivity Factors: What factors (e.g., work hours, overtime, training hours) most influence employee productivity, and how do they vary across departments?**

It was assumed that the metric for productivity was based on performance scores. We see that the performance scores across the departments range from 2.98 to 3.02 this would indicate there are no significant differences in performance scores across the departments. The highest average performing department was Engineering at 3.02 rating. An interesting insight is that within the Engineer job title which has PHD as the education level has an average of 2.91 rating. This is a very low rating, and this could be attributed to the fact that they earn the least in the Engineering job title.

A correlation analysis was done in Power BI and it revealed that salary has the biggest relation to Performance /productivity of an employee on a departmental level. The factors and their correlation coefficient can be depicted in the dashboard. The weakest impact is the training hours.

# Recommendations

Power BI engine should not be used for denormalizing the data this should be done on a database level and Power BI should just bring in normalized tables.

An Analysis on Male vs Female can be done to see the satisfaction and performance levels to compare and see if there are significant variations.

Based on the data given a predicative burnout index could be created since sick leave and resignation data is available this may help to understand and optimize employees mental and physical health.