**Business Analytics Project** 

BUS 2110 – Management Information Systems

Gisselle Garcia-Jose

Human Resources Dataset

# Introduction

The dataset given to me as your consultant to analyze was a human resources dataset from an unnamed organization with given instructions to analyze the dataset and provide feedback on various Data Driven Opportunities as well as through the creation of data driven opportunities. The dataset given to our consulting company included various columns (including company, company code, employee, employee status, employee status descriptions, department, department name, job code, job code description, date hired, salary class, exempt employee, auto deposit description, auto deposit, termination date, creation date, deceased, work state, tax state, tipped, death date, salary, last evaluation score, active status, length of employment and active employee.). In the Appendix, a Data Dictionary (giving detailed definitions of the various fields provided in the dataset) outlines the provided data in more detail. The dataset consisted of X number of individuals as where the columns provide more information about the individual within the company.

## List of Data Driven Opportunities

The Human Resources consulting team was given three specific Data Driven Opportunities to answer as a part of the project and was asked to create two additional Data Driven Opportunities to research. Given the instructions of the project, the following list of question regarding the Data Driven Opportunities will be attempted to be answered as a part of the project:

- What Departments have the highest and lowest average evaluation scores? (Provided)
- Which Jobs have the highest and lowest average salaries? (Provided)
- Excluding Terminated Employees, which Department has the employees with the longest and shortest average Length of Employment? (Provided)
- Which departments have the highest and lowest average salaries? (Supplemental Data Driven Opportunity)
- Which work state has the highest average evaluation scores? (Supplemental Data Driven Opportunity)

### List of Required Fields

Based on these Data Driven Opportunities, I identified a list of all the fields from the Data Dictionary which would be needed to complete this research as well as any needed calculated fields. The below listed outlines the required fields:

- Department Name
- Job Code
- Job Code Description
- Employee Status Description
- Work State
- Last Evaluation Score
- Active Status?
- Length of Employment
- Active Employee

This is the output I would need from the Cleaning Process to be able to perform the requested analysis.

## **Review of the Cleaning Process**

Using the list above, I determined that the following fields from the Main dataset would not be needed for the requested analysis and could be removed:

- Company code
- Employee status
- Exempt employee
- Auto deposit description
- Auto deposit
- Creation date
- Deceased
- Tax state
- Tipped
- Death date

Next, I used if functions in Excel to add the following fields to the dataset:

- Company Code
- Salary class
- Exempt employee
- Auto direct deposit description
- Tipped

Lastly, I looked for suspicious data and found the following:

- Suspicious Data #1 ---- Removed the 1800-01-01 data by using the filter in the term\_date column. I removed the 1800-01-01 data because it was just the default null value instead of leaving the slot blank. It was not affecting the data so, I decided to remove it through filtering.
- Suspicious Data #2 ---- Removed blanks from the Auto\_deposit column. As I was analyzing the data, I noticed that some rows had many blanks so I chose to remove them in order to condense the data some rather than just keep it in there. Again, due to it not really providing any information, or more or so lacking information. I chose to remove them.

At this point, a final cleaned dataset was saved and was used as input to the next step in the process.

### Analysis, Findings, Recommendations

The first Data Driven Opportunity I was assigned was:

• What Departments have the highest and lowest average evaluation scores?

To research this data driven opportunity, I chose to create two separate Bar Charts due to convenience and readability.





As you can see, the department with the lowest average evaluation scores is FC Ranch Pool and the department with the highest average evaluation scores is Utilities Lake. This being said, it would be best to focus on the departments that have a lower average evaluation score.

Moving on to the next question.

• Which Jobs have the highest and lowest average salaries? (Provided)



<Jobs with Hightest Average Salaries>

Avg. Salary 🗐



From the analysis and research provided above, there appears to be a wide range of jobs that have the lowest and highest average of salaries. Most of the positions shown in both data displays are fairly close in value which I was surprised by.

• Excluding Terminated Employees, which Department has the employees with the longest and shortest average Length of Employment? (Provided)



< Excluding Terminated Employees, Department with employees with the longest average Length of Employment >



MIS is the department with employees with the longest average length of employment and it appears that the Depot Gift Shop is the department with employees with the shortest average length of employment. The difference between the two averages is highly drastic. It is very interesting information to analyze.

• Which departments have the highest and lowest average salaries? (Supplemental Data Driven Opportunity)



<Departments with highest salaries>

#### <Departments with lowest salaries>



Lastly, the second supplemental data driven opportunity is the following:

### • Which work state has the highest evaluation scores? (Supplemental Data Driven Opportunity)

<Work State with hightest average evaluation scores>



<Work State with hightest average evaluation scores>



For this second opportunity, I chose to display the information not only using a tree map but also, by using a map because the data points appeared the most interesting to me. With the tree map it is clearly visible that New Hampshire is the work state with the highest evaluation score. However, even though on the second map it is hard to tell, I liked to see both charts side by side in order to provide a better comparison. It is interesting to see how different data displays still portray the same information and story.

This business analytics project provided a lot of valuable insight to the overall approach and delivery of the information. Many columns of data were used repeatedly however still analyzed in different ways according to the question at hand.

### Appendix

Data Set	
COMPANY NAME	Name given to each department of the company
EMPLOYEE	Unique number assigned to each employee purchased as a part
	of this order
EMPLOYEE STATUS	Displays status of employees in dataset. TF meaning termination
	final and AF meaning active full time
EMPLOYEE STATUS DESCRIPTION	Written out description of EMPLOYEE STATUS
DEPARTMENT	Department in company where employee is employed
JOB CODE	Unique code for each job within department

Below is the Data Dictionary outlining all the fields we were given to analyze:

JOB CODE DESCRIPTION	Written out description of JOB CODE
DATE HIRED	Date that the employee was hired
SALARY CLASS	Displays whether an employee is paid "Hourly" or "Salaried"
EXEMPT CODE	Displays whether an employee "Uses Time Clock" or "Does not use Time Clock"
AUTO DEPOSIT	Displays whether an employee's direct deposit is "Active" or "Inactive"
TERMINATION DATE	Date that employee was termination. Displays 1800-01-01 if still active employee.
CREATION DATE	Date when employee account was made
DECEASED	Displays whether the employee is deceased or not
WORK STATE	Displays the state in which the employee is working
TAX STATE	Displays the state in which the employee is taxed.
TIPPED	Whether and employee is eligible or not for being tipped
DEATH DATE	Displays date of employee's death, if not dead displays
SALARY	Employee's salary when employed
LAST EVALUATION SCORE	Displays grade of last evaluation by Poor. Below Average.
	Average, Above Average, Excellent
LENGTH OF EMPLOYMENT	Length of employee's employment in days
ACTIVE EMPLOYEE	Displays whether an employee is active or has been terminated