

Data Analysis Project

Data Analyst: (Carlos Silverio)

Client/Sponsor: HR Department, AcmeTech Solutions

Purpose: *Write a brief description of why this project is happening below. Why is this project happening? What are the goals?*

AcmeTech Solutions, a large company with approximately 4,000 employees, faces a 26% annual employee turnover rate, which is concerning for several reasons.

1. This turnover leads to project delays, negatively impacting the company's reputation with consumers and partners.
2. Additionally, maintaining a substantial recruitment department is necessary to replace departing employees, and new hires often require training and adjustment time.

The management has enlisted an HR analytics firm, where I am a key analyst, to identify the primary factors contributing to attrition and recommend workplace changes to reduce it. The goal of this project is to determine what are the main reasons causing attrition in the firm, so the company can make changes accordingly. In other words, they want to know what changes they should make to their workplace, to get most of their employees to stay. Also, they want to know which of these variables is most important and needs to be addressed right away.

Project Goals:

- Identify the primary factors contributing to attrition within AcmeTech Solutions.
- Recommend workplace changes to reduce attrition.
- Prioritize the factors that have the most significant impact on attrition.

Success Metrics: The project will be considered successful if it:

- Identifies the top three factors contributing to attrition.
- Recommends actionable changes based on data analysis.
- Reduces attrition by at least 10% within the next year.

Assumptions and Constraints:

- The provided data is accurate and complete.
- Access to necessary data sources is granted.
- The project will focus on variables related to employee satisfaction and performance.

Risk Assessment: Potential risks include:

- Data quality issues.
- Delays in data collection.
- Unexpected changes in project requirements.

Collaboration and Communication Plan:

- Weekly project updates will be provided to stakeholders.
- Feedback will be actively sought from stakeholders throughout the project.

Data Privacy and Security: All data used will be anonymized and handled in compliance with relevant data privacy regulations.

Visualizations and Reporting:

- Tableau will be used for data visualizations.
- Microsoft PowerPoint will be used for presentations.

Quality Assurance: Peer reviews of code and statistical methods will be conducted to validate findings.

Mid-Project Report: The mid-project report will include:

- Preliminary findings and insights.
- Any challenges encountered and their resolutions.
- Adjustments to the project plan.

Data Dictionary: A data dictionary will be provided to aid in understanding the dataset.

Feedback Loop: Stakeholder feedback will be incorporated into the project at regular intervals.

Conclusion and Next Steps: After presenting findings and recommendations, the project will move to the implementation phase, focusing on making workplace changes based on the analysis.

To solve this problem, I must answer the following questions that came up in a meeting with my stakeholders:

1. How many employees does the company have today?
2. How many employees left the company last year?
3. What is the percentage of employees who left the company last year?
4. How many employees with low job satisfaction left the company last year?

5. What is the percentage of employees with low job satisfaction who left the company?
6. How many employees with low job environment satisfaction left the company last year?
7. What is the percentage of employees with low job environment satisfaction who left the company?
8. How many employees with low work-life balance satisfaction left the company last year?
9. What is the percentage of employees with low work-life balance satisfaction who left the company?
10. How many employees with low job involvement left the company last year?
11. What is the percentage of employees with low job involvement who left the company?
12. How many men and women left the company last year?
13. What is the percentage of men who left the company last year?
14. What is the percentage of women who left the company last year?
15. How many employees left the company for each department?
16. How many employees left the company for each job role?
17. What is the average duration of an employee's tenure?
18. What is the average duration of employment for employees in each department?
19. How many employees had salaries below the department average and left the company?
20. What is the percentage of employees who left the company and earned below the department average?
21. What is the average number of years since the last promotion for employees who left?
22. What is the average number of years since the last promotion for employees who stayed?
23. What is the average tenure (total working years) of employees who left?
24. What is the average tenure (total working years) of employees who stayed?

Scope / Major Project Activities:

What are the major parts of this project? List out the high-level steps, activities, or stages of the project, and give a brief description for each.

Activity	Description
Understand the problem:	Clearly define and understand the problem. Get in touch with stakeholders and understand what they are looking out for from this project. Ask as many questions as you can. Come up with questions that can help you solve the problem.
Collect the data:	Now that you know the problem, and your stakeholders and you are on the same page, is time to collect data that will help you solve it. Once you collect the data, review it. Understand the data set and see how the data set can help you solve the problem at hand. Come up with 10-5 questions that might help you get to the solution of the problem.
Data Preprocessing: involves tasks such as data cleaning, handling missing values, outlier treatment, join tables etc. get rid of unnecessary columns etc.	Cleanse and preprocess the employee datasets to handle missing values, outliers, and inconsistencies, ensuring data quality and integrity for accurate analysis. In this step, I will also join tables since I am working with three tables. I will also exclude any unnecessary column.
Exploratory Data Analysis (EDA): It is about making sense or understanding the data. Visually review the data, bar graph, scatterplots, correlation, boxplot etc.	Perform EDA to gain initial insights into employee's behaviors and their relationship with attrition. Write queries to answer the questions you and your stakeholders came up at the beginning of the project.
Share your results (visualization):	Create a dashboard that shows what has caused employees to leave the company. Show relationships, trends, any information that might help your stakeholders understand the reason why employees are leaving the company.
Recommendation	Create a power point presentation with your findings and recommendations to present to your stakeholder's from HR department. They are not very technical, so make the presentation as easy to understand as possible.

This project does not include:

Specify the things that this project isn't responsible for doing (out of scope). For instance, "this project does not involve a summation of 2019 data analysis"

- This project will focus on the variables that are related to the satisfaction and performance of the employee in the company. Other variables won't be considering for the result of this analysis. I will do some analysis using those variables for practice purpose for I will only focus on satisfaction and performance variables.

Deliverables:

A specific list of things that your project will deliver.

Deliverable	Description/ Details
Create a scope of work	Create a detailed outline showing how are you going to tackle this project, including deliverables, main goals, due dates etc.
Create a mid-project report.	Create a mid-project report to let your clients know how the project is going, if something has raised, and you need something from them etc.
Tableau Dashboard for HR.	Create a tableau dashboard for your clients.
PowerPoint presentation for HR.	Create a PowerPoint presentation for your clients with recommendations on what to do next.
Upload your work to github / medium	Once your work is over, upload it to GitHub and medium.

Schedule Overview / Major Milestones:

The expected schedule for the project. This can be defined by milestones (e.g. "all data is cleaned and processed"), periods of time ("Week 1 / Week 2"), or other ways based on the needs of the project.

Milestone	Expected Completion Date	Description/Details
ASK face of the process	9/13/2023 - 9/14/2023	define the problem, understand what your responsibilities are, understand the stakeholders' expectations, finish the scope of work, Check the credibility of the datasets. Create questions for the analysis project.
Preprocess the data	9/14/2023 - 9/21/2023	Handle missing values, duplicates, inconsistencies, remove unnecessary columns, merged the datasets etc.

Analyze the data	9/18/2023 -9/22/2023	Write queries to answer the questions related to the project.
Dashboard creation	9/23/2023 - 9/24/2023	Create a dashboard that displays the main reasons why the employees are leaving the company.
Present results	9/25/2023	Create a power point presentation that summarizes your findings. Include graphs, and support material.

*Estimated date for completion:

I started the project on 9/12/2023. If all goes well and I have everything I need, the project will be ready by 9/26/2023.