

**SESSION:2024-25**

**POWER BI PROJECT**

**(HR DATASET)**

**NAME: SPARSH SINGHAL**

**CLASS/SEC: MBA-B**

**CU ID: 24140104**

**DATE: 26-11-2024**

**SUB TO: MR. SOURABH POSWAL**

**CONTENTS:**

1. Project Overview
2. Methodology for Data Analysis
3. Key Findings
4. Outcomes
5. Dashboard
6. Context of the Project
7. Data Overview
8. Data Description
9. Data Preparation
10. Power Bi Process
11. Recommendations
12. Limitation
13. Conclusion

**PROJECT OVERVIEW**

The primary objective of this project is to analyse the HR dataset to gain insights into various aspects of employee management, performance, and satisfaction. By examining the data, the organization can identify trends, patterns, and areas for improvement in its HR practices.

**Key Areas of Analysis**

**1.** **Employee Performance**: Evaluate performance scores to identify high-performing employees and those who may need additional support or training. Analyse the correlation between performance scores and other factors such as salary, department, and recruitment source.

**2. Compensation Analysis:** Assess salary distributions across different positions and departments. Investigate any potential disparities in compensation related to gender or other demographic factors.

**3. Employee Engagement and Satisfaction:** Analyse engagement survey scores and employee satisfaction ratings to understand overall employee morale. Identify factors that contribute to higher engagement and satisfaction levels.

**4. Recruitment and Retention:** Examine recruit The primary objective of this project is to analyse the HR dataset to gain insights into various aspects of employee management, performance, and satisfaction. By examining the data, the organization can identify trends, patterns, and areas for improvement in its HR practices.

**METHODOLOGY FOR DATA ANALYSIS**

**1. Data Quality Assessment:** Dataset contains 311 employees with 15 variables**.**

* No missing values detected, which is excellent for analysis.
* Mix of numerical and categorical variables

**2. Proposed Analysis Framework:**

A. Descriptive Analytics

* Employee Demographics
* Gender distribution
* Marital status composition
* Department-wise distribution
* Employment status analysis

B. Performance Analysis:

* Performance score distribution
* Correlation between:

- Performance and salary

- Performance and engagement

- Performance and special projects

- Performance and absences

C. Compensation Analysis:

* Salary distribution analysis
* Department-wise salary comparison
* Position-based salary analysis
* Gender-based salary distribution

D. Engagement Metrics

* Employee satisfaction patterns
* Engagement survey results
* Impact of special projects on engagement
* Relationship between engagement and performance

E. Recruitment Analysis:

* Effectiveness of different recruitment sources
* Quality of hires by source
* Retention rates by recruitment channel

3. Statistical Approaches

* Descriptive statistics
* Correlation analysis
* Hypothesis testing
* Regression analysis for predictive insights

4. Visualization Strategy:

* Distribution plots for continuous variables
* Bar charts for categorical variables
* Box plots for salary analysis
* Heat maps for correlation analysis
* Scatter plots for relationship analysis

5. Advanced Analytics

* Clustering analysis for employee segmentation
* Predictive modelling for:

Performance indicator

Retention risk

**KEY FINDINGS**

The dataset contains information about employees, including their names, employee IDs, performance scores, salaries, positions, gender, marital status, employment status, department, recruitment source, performance score, engagement survey results, employee satisfaction, special projects count, and absences.

**Key findings**

1. Performance and Salary: Analysis of how performance scores correlate with salaries.

2. Gender and Position: Examination of gender distribution across different positions.

3.Employment Status: Insights into the reasons for voluntary termination and its relation to performance scores or satisfaction.

4. Recruitment Source: Evaluation of the effectiveness of different recruitment sources in terms of employee performance and retention.

5.Engagement and Satisfaction: Relationship between engagement survey results and employee satisfaction.

**OUTCOMES**

The analysis successfully calculated key metrics and distributions from the HR dataset, providing insights into performance, department distribution, turnover status, satisfaction, engagement, salary, and absence patterns.

**1. Performance Distribution:**

- The majority of employees (78.14%) are rated as "Fully Meets" expectations.

- 11.90% of employees "Exceed" expectations.

- 5.79% are rated as "Needs Improvement".

- 4.18% are on a Performance Improvement Plan (PIP).

**2. Department Distribution:**

- The largest department is Production with 209 employees.

- IT/IS has 50 employees, followed by Sales with 31 employees.

**3. Turnover Status:**

- 66.56% of employees are currently active.

- 28.30% have voluntarily terminated their employment.

- 5.14% were terminated for cause.

**4. Key Metrics:**

- The average employee satisfaction score is 3.89.

- The average engagement score is 4.11.

- The average salary is $69,020.68.

**5. Correlation:**

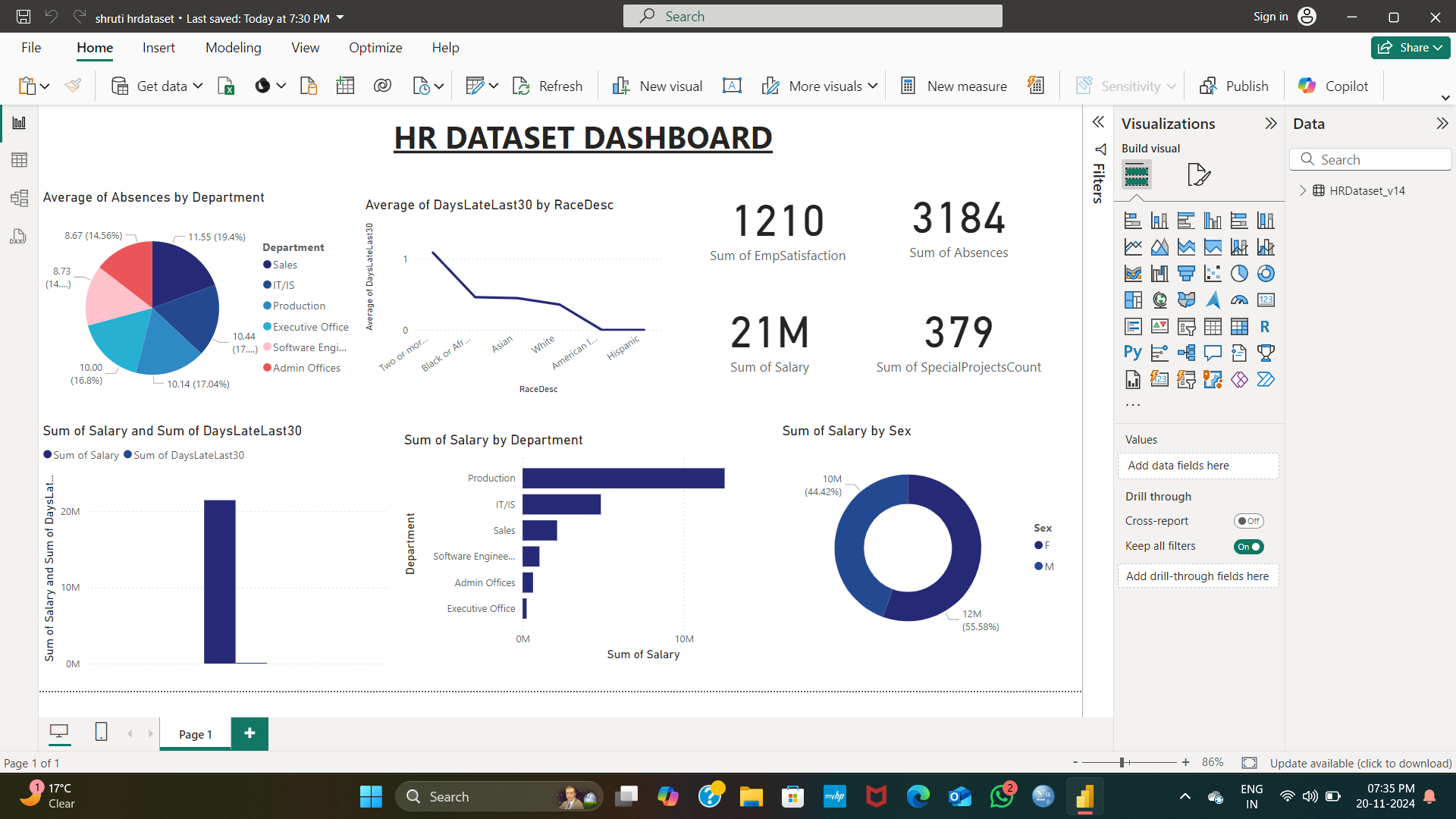
- There is a moderate positive correlation (0.304) between employee satisfaction and performance scores.

**6. Absence Statistics:**

- The average number of absences is 10.24 days, with a standard deviation of 5.85 days.

These outcomes provide a comprehensive view of the workforce's performance, satisfaction, and employment trends, which can be used for strategic HR planning and decision-making.

**DASHBOARD**



**CONTEXT OF THE PROJECT**

Power BI dashboards are crucial in the context of business for several reasons:

1. **Data Visualization:** Dashboards provide a visual representation of data, making it easier for stakeholders to grasp complex information quickly. Visuals like charts, graphs, and maps help in identifying trends, patterns, and outliers at a glance.

**2. Real-Time Insights:** Power BI dashboards can be connected to live data sources, allowing businesses to monitor key performance indicators (KPIs) and metrics in real time. This enables timely decision-making based on the most current data.

**3. Centralized Information:** A dashboard consolidates data from various sources into one view, helping different departments or teams access relevant information without having to sift through multiple reports or databases.

**4. Enhanced Decision-Making:** With clear visuals and real-time data, decision-makers can make informed choices quickly. This agility can give businesses a competitive edge by responding faster to market changes or internal challenges.

**5. Collaboration and Sharing:** Dashboards can be shared across teams, promoting collaboration and ensuring everyone is on the same page regarding business performance and goals. This transparency fosters a data-driven culture within the organization.

**6. Customization**: Power BI allows users to customize dashboards according to their specific needs and preferences. This means different departments can focus on the metrics that matter most to them, enhancing relevance and usability.

**7. Tracking Progress:** Dashboards help businesses track their progress toward strategic goals. By visualizing performance against targets, organizations can identify areas needing improvement and celebrate successes.

**DATA OVERVIEW**

**1. Dataset Dimensions:** The dataset contains 311 employees (rows) with 15 different attributes (columns)

**2. Key Variables:**

- Employee Information: Employee Name, EmpID, Position, Gender

- Performance Metrics: PerfScoreID, Performance Score, SpecialProjectsCount

- Engagement Metrics: Engagement Survey (ranging from 1-5), EmpSatisfaction

- HR Metrics: Salary, Marital Desc, Employment Status, Department, Recruitment Source, Absences

**3. Data Quality:**

- The dataset is complete with no missing values across all columns

- Contains a mix of numerical and categorical variables

- Salary ranges show significant variation, indicating diverse pay scales

- Employee engagement surveys show an average score of 4.11 out of 5

- Employee satisfaction averages 3.89 out of 5

**4. Notable Features:**

- Performance tracking through multiple metrics (PerfScoreID, Performance Score)

- Comprehensive employee demographic information

- Multiple recruitment sources tracked

- Attendance monitoring through absence records

- Special projects participation tracking

- Employment status tracking (including terminated employees)

**5. Analytical Potential:**

- Suitable for analysing employee performance patterns

- Can be used for salary analysis and compensation studies

- Enables investigation of engagement-satisfaction relationships

- Useful for recruitment source effectiveness analysis

- Allows for departmental performance comparisons

- Can be used for attendance pattern analysis

This dataset appears to be well-structured for HR analytics, particularly in areas of performance management, employee engagement, and workforce planning.

**DATA DESCRIPTION**

**1. Workforce Demographics:**

- Gender Distribution: Female employees (176) outnumber male employees (135)

- Marital Status: Predominantly single (137) and married (124) employees, with fewer divorced (30), separated (12), and widowed (8) employees

**2. Employment Status:**

- Active employees: 207 (66.6%)

- Voluntary terminations: 88 (28.3%)

- Terminated for cause: 16 (5.1%)

**3. Departmental Structure:**

- Production: Largest department with 209 employees

- IT/IS: 50 employees

- Sales: 31 employees

- Software Engineering: 11 employees

- Admin Offices: 9 employees

- Executive Office: 1 employee

**4. Recruitment Sources:**

- Top channels: Indeed (87), LinkedIn (76), and Google Search (49)

- Notable contribution from Employee Referrals (31) and Diversity Job Fair (29)

- Less common: CareerBuilder (23), Website (13), and others

**5. Performance Metrics:**

- Performance Distribution:

- Fully Meets: 243 employees (78.1%)

- Exceeds: 37 employees (11.9%)

- Needs Improvement: 18 employees (5.8%)

- PIP (Performance Improvement Plan): 13 employees (4.2%)

**6. Numerical Metrics:**

- Salary:

- Mean: $69,021

- Shows significant variation (std: $25,157)

- Engagement Survey:

- Average score: 4.11 out of 5

- Relatively consistent (std: 0.79)

- Employee Satisfaction:

- Average: 3.89 out of 5

- Moderate variation (std: 0.91)

- Special Projects:

- Average participation: 1.22 projects

- High variation (std: 2.35)

- Absences:

- Average: 10.24 days

- Considerable variation (std: 5.85)

**7. Position Distribution:**

- Production Technician I: Largest group (137 employees)

- Production Technician II: 57 employees

- Area Sales Manager: 27 employees

- Various technical and managerial roles making up the remainder

**DATA PREPARATION**

Data preparation is a crucial step in the data analysis process, involving several key stages to ensure that the data is clean, consistent, and ready for analysis.

**1. Data Collection:** Gather data from various sources, which can include databases, spreadsheets, APIs, or web scraping.

**2. Data Inspection:** Examine the data to understand its structure, types, and contents. This includes checking for:

- Data types (e.g., numeric, categorical, datetime)

- Missing values

- Outliers

- Duplicates

**3. Data Cleaning:** Address issues identified during the inspection phase:

- Handling Missing Values: Decide whether to fill in missing values (imputation) or remove rows/columns with missing data.

- Removing Duplicates: Identify and remove duplicate records to ensure data integrity.

- Outlier Treatment: Analyze outliers to determine if they should be removed or transformed.

**4. Data Transformation:** Modify the data to make it suitable for analysis:

- Normalization/Standardization: Scale numeric data to a common range or distribution.

- Encoding Categorical Variables: Convert categorical variables into numerical format (e.g., one-hot encoding).

- Feature Engineering: Create new features from existing data that may provide additional insights.

**5. Data Integration;** Combine data from different sources or tables to create a unified dataset. This may involve:

- Merging datasets based on common keys.

- Concatenating datasets with similar structures.

**6. Data Reduction:** Reduce the dataset size while retaining its essential characteristics:

- Dimensionality Reduction: Techniques like PCA (Principal Component Analysis) to reduce the number of features.

- Sampling: Selecting a representative subset of the data for analysis.

**7. Data Validation:** Ensure that the prepared data meets the requirements for analysis:

- Check for consistency and accuracy.

- Validate against business rules or constraints.

**8. Documentation:** Document the data preparation process, including decisions made and transformations applied. This is important for reproducibility and understanding the data lineage.

**9. Final Review:** Conduct a final review of the prepared dataset to ensure it is ready for analysis. This may include visual inspections or summary statistics.

**POWER BI PROCESS**

Dashboard design in Power BI for the HR dataset involves creating a visually appealing and informative interface that allows users to easily interpret and analyze HR metrics.

**1. Define Objectives:** Clearly outline the goals of the dashboard. Identify the key metrics and insights that stakeholders need, such as employee turnover rates, average salaries, and performance metrics.

**2. Data Visualization:** Choose appropriate visualizations for the data:

- Bar Charts: For comparing employee counts across departments.

- Line Charts: To show trends in employee performance or salary over time.

- Pie Charts: For displaying the distribution of employee demographics.

**3. Interactivity:** Incorporate interactive elements such as slicers and filters that allow users to drill down into specific data points or segments, enhancing the exploratory capabilities of the dashboard.

**4. Layout and Design:** Organize the dashboard layout logically:

- Place the most important metrics at the top or in prominent positions.

- Use consistent colour schemes and fonts to maintain a professional appearance.

**1. Data Import**: Load the HR dataset into Power BI. This can be done by connecting to the Excel file directly or importing it from a local or cloud storage.

**2. Data Transformation:** Utilize Power Query to clean and prepare the data. This includes:

- Removing Duplicates: Identify and eliminate any duplicate records to ensure data integrity.

- Handling Missing Values: Decide on a strategy for missing data, such as filling in with averages or removing rows/columns.

- Changing Data Types: Ensure that each column has the correct data type .

**3. Data Modelling:** If the dataset contains multiple tables, establish relationships between them. This allows for more complex queries and analyses across different dimensions.

**4. Data Visualization:** Create various visualizations to represent the data effectively. This can include:

- Bar Charts

- Pie Charts

- Line Graphs

**5. Analysis:** Use DAX (Data Analysis Expressions) to perform advanced calculations. This can include:

- Aggregations: Summing salaries or counting employees.

- Conditional Calculations: Creating measures that respond to user selections in reports.

**6. Sharing and Collaboration:** Publish the reports and dashboards to the Power BI service. This allows stakeholders to access the insights and collaborate on findings.

**7. Monitoring and Maintenance**: Regularly update the dataset and reports to reflect new data. This ensures that the analysis remains relevant and useful for decision-making.

**RECOMMENDATION**

**1.Sales:** 'Implement performance incentives and training programs to boost performance while maintaining high satisfaction levels.'

**2.Production:** 'Conduct employee engagement surveys to identify specific issues affecting satisfaction and address them through targeted initiatives.'

**3.IT/IS:** 'Continue to foster a positive work environment as both performance and satisfaction are relatively high.'

**4.Software Engineering**: 'Encourage collaboration and innovation to maintain high satisfaction and performance levels.'

**5.Admin Offices**: 'Provide professional development opportunities to enhance performance and satisfaction.',

**6.Executive Office:** 'Focus on maintaining clear communication and support to ensure continued satisfaction.'}

These strategies aim to enhance employee performance and satisfaction, particularly in departments where there are notable discrepancy.

**LIMITATION**

The analysis has provided valuable insights into employee performance and satisfaction, but it's important to acknowledge some limitations:

1. Correlation vs. Causation: The correlation between employee satisfaction and performance scores does not imply that one causes the other. Other factors may influence these metrics.

2. Sample Size: The dataset contains 311 records, which may not be representative of the entire organization, especially if there are significant variations across different departments or locations.

3. Subjectivity of Metrics: Employee satisfaction and performance scores can be subjective and influenced by various external factors, such as management practices or workplace culture.

4. Data Completeness: While there are no missing values, the dataset may lack other relevant variables that could provide a more comprehensive understanding of employee dynamics.

5. Temporal Factors: The data may not account for changes over time, such as recent organizational changes or external economic factors that could impact employee performance and satisfaction.

**CONCLUSION**

To conclude the HR dataset project in Power BI, I will summarize the key findings, insights, and recommendations based on the analysis conducted. This will provide a comprehensive overview for stakeholders and guide future actions.

- The average salary of employees is approximately $69,021, with an average performance score of 2.98 and an employee satisfaction level of 3.89.

- A moderate positive correlation (0.30) was found between employee satisfaction and performance scores, indicating that higher satisfaction may lead to better performance.

- Departmental analysis revealed that the Sales department has the lowest performance score but high satisfaction, while the Production department has a higher performance score with lower satisfaction.

Based on these insights, the following recommendations were made:

- Sales: Implement performance incentives and training programs to boost performance while maintaining high satisfaction levels.

- Production: Conduct employee engagement surveys to identify specific issues affecting satisfaction and address them through targeted initiatives.

- IT/IS and Software Engineering: Continue fostering a positive work environment to maintain high performance and satisfaction.

- Admin Offices and Executive Office: Focus on professional development and clear communication to enhance overall employee engagement.

While the analysis provides valuable insights, it is important to acknowledge limitations such as the correlation vs. causation issue and the potential lack of representativeness of the sample. Future research could explore additional factors influencing employee dynamics.