# EMPLOYEE DATA ANALYSIS USING EXCEL

Name:D.Jothi

Register number :asunm110312201347

Department:b.com(commerce)

College:D.R.B.C.C.C Hindu college

# AGENDA

✔Problem statement

✔project overview

✔end users

✔our solution and proposition

✔dataset to description

✔modelling approach

✔Result and discussion

✔conclusion

# PROBLEM STATEMENT

***Turnover Comparison***: How does the turnover differ between male and female employees across different joining years?

***Trend Over Time***: Is there a trend in turnover amounts over the years, and does it differ by gender?

***Gender Gap:*** What is the average turnover for male versus female employees, and does the data suggest any significant gender-based differences?

***Year-wise Analysis:*** Are there specific years where the turnover significantly differs, and what might be contributing to these differences?

# PROJECT OVERVIEW

**2012:**
**Male:** 3 employees, Turnover Range: 85,879 - 105,469
**Female:** 1 employee, Turnover: 88,361
**2013:**
**Male:** 5 employees, Turnover Range: 57,002 - 118,976
**Female:** None
**2014:**
**Male:** 4 employees, Turnover Range: 68,981 - 114,425
**Female:** 2 employees, Turnover Range: 42,314 - 74,279
**2016:**
**Male:** 6 employees, Turnover Range: 37,902 - 52,964
**Female:** 4 employees, Turnover Range: 52,749 - 69,913
**2017:**
**Male:** 10 employees, Turnover Range: 43,329 - 90,884
**Female:** 7 employees, Turnover Range: 50,449 - 104,335
**2018:**
**Male:** 3 employees, Turnover Range: 69,163 - 113,616
**Female:** 1 employee, Turnover: 110,906

# END USERS

**2012:** Highest turnover in this year is 105,469 (Male).

**2013**: The turnover is more concentrated among males, with a high variance**.**

**2014**: Both males and females have a significant turnover, with females showing more consistent lower turnover**.**

**2016:** Lower turnover for males compared to females in this year.

**2017:** High turnover across both genders, with some very high turnovers for females.

**2018:** Significant turnover for both genders, with one male (113,616) and one female (110,906) showing the highest turnovers.

# solution

Gender-Based Analysis:

AverageTurnover by Gender:

Calculate the average turnover for male and female employees to identify any significant gender disparities.

Turnover Variability:

Analyze the range and variability of turnover amounts within each gender to understand if one gender experiences more turnover volatility than the other.

Year-Based Analysis:

Yearly Turnover Trends:

Track turnover trends over the years to see if newer employees have different turnover rates compared to longer-tenured employees. Focus on identifying if certain years (e.g., 2017) show different patterns.

Cohort Analysis:

Group employees by their joining year and analyze turnover within these cohorts. This can help in understanding whether specific cohorts are experiencing higher turnover and why.

Statistical Insights:

T-Test/ANOVA:

Conduct statistical tests to determine if the differences in turnover between genders or across joining years are statistically significant.Correlation Analysis:Assess the correlation between joining year and turnover, and whether it differs by gender. This can help identify whether newer employees, regardless of gender, have a higher turnover.

Predictive Modeling:

Regression Analysis:Use regression models to predict turnover based on gender and joining year. This will help in identifying key factors that influence turnover rates and in forecasting future turnover tr

# proposition

Address Gender Disparities:If the analysis reveals significant gender disparities in turnover, propose interventions such as regular pay equity audits and targeted support programs to address underlying issues.Enhanced Retention Strategies:For cohorts with higher turnover (e.g., 2017 joiners), suggest targeted retention strategies like mentorship programs, career development opportunities, or enhanced benefits that cater specifically to the needs of these groups.Diversity and Inclusion Initiatives:Strengthen diversity and inclusion initiatives if female employees show significantly different turnover rates. This could include gender-specific support systems, flexible work arrangements, and leadership development programs.Continuous Monitoring and Reporting:Implement a system for continuous monitoring of turnover rates by gender and joining year, with regular reports to management. This will ensure that any emerging issues are identified and addressed promptly.Employee Feedback Mechanism:Establish or enhance existing feedback mechanisms to directly gather insights from employees, particularly from those cohorts or gender groups with higher turnover rates. This qualitative data can provide deeper understanding and actionable insights.

# DATASET TO DESCRIPTION

Joining Year:The year in which the employee joined the company. This ranges from 2012 to 2018.Gender:The gender of the employee, categorized as either "MALE" or "FEMALE."Turnover:The monetary value of turnover associated with each employee. This is a continuous numerical variable that represents the turnover amount in some currency (e.g., dollars).

Joining Year Distribution:2012: 5 records2013: 6 records2014: 6 records2016: 9 records2017: 14 records (the highest concentration)2018: 5 recordsGender Distribution:MALE: 26 recordsFEMALE: 19 records

# modelling approach

DATA PREPARATIONDATA CLEANING: CHECK FOR MISSING OR INCORRECT VALUES.ENCODING: CONVERT CATEGORICAL VARIABLES LIKE GENDER INTO NUMERICAL FORMAT (E.G., MALE = 0, FEMALE = 1).FEATURE ENGINEERING: CONSIDER ADDING NEW FEATURES SUCH AS THE DIFFERENCE BETWEEN THE JOINING YEAR AND THE CURRENT YEAR TO REPRESENT EXPERIENCE.2. EXPLORATORY DATA ANALYSIS (EDA)DESCRIPTIVE STATISTICS: COMPUTE MEAN, MEDIAN, AND STANDARD DEVIATION FOR TURNOVER ACROSS GENDERS AND JOINING YEARS.CORRELATION ANALYSIS: CHECK CORRELATIONS BETWEEN THE FEATURES AND TURNOVER.VISUALIZATIONS: PLOT HISTOGRAMS, BOX PLOTS, AND SCATTER PLOTS TO OBSERVE RELATIONSHIPS AND DISTRIBUTIONS.3. MODELING APPROACHESLINEAR REGRESSION: USE LINEAR REGRESSION IF THE RELATIONSHIP BETWEEN FEATURES AND TURNOVER IS EXPECTED TO BE LINEAR.DECISION TREES/RANDOM FOREST: SUITABLE FOR CAPTURING NON-LINEAR RELATIONSHIPS BETWEEN FEATURES.GRADIENT BOOSTING MACHINES (GBM): USEFUL FOR MORE COMPLEX RELATIONSHIPS AND HIGHER PREDICTIVE POWER.NEURAL NETWORKS: IF THE DATASET WERE LARGER, NEURAL NETWORKS COULD BE CONSIDERED FOR CAPTURING COMPLEX PATTERNS.4. MODEL EVALUATIONTRAIN/TEST SPLIT: SPLIT THE DATA INTO TRAINING AND TESTING SETS (E.G., 80% TRAINING, 20% TESTING).METRICS: USE R-SQUARED, MEAN ABSOLUTE ERROR (MAE), AND ROOT MEAN SQUARE ERROR (RMSE) FOR EVALUATING REGRESSION MODELS.CROSS-VALIDATION: APPLY CROSS-VALIDATION TO ENSURE THE MODEL'S ROBUSTNESS.

# RESULT
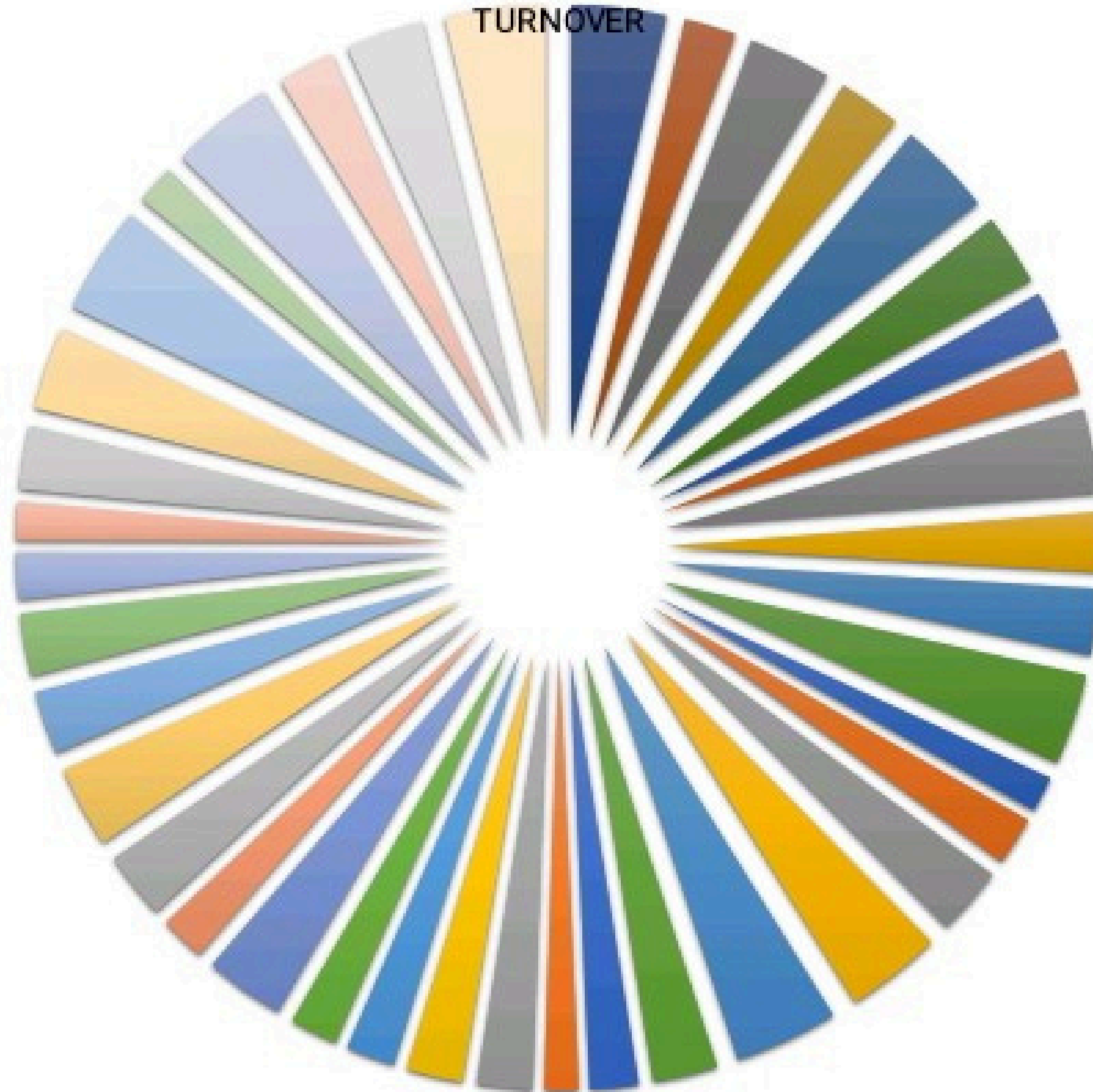


TURNOVER

# conclusion

While the data provides a snapshot of turnover by gender and joining year, the variance within the groups suggests that other factors are likely at play in determining turnover. Gender alone does not seem to be a significant determinant, but the joining year shows a trend where earlier joiners tend to have higher turnovers. Further analysis with additional data could provide more insights into the factors influencing turnover.

● ● ●

# THANK YOU