HR Analytics – Employee Attrition
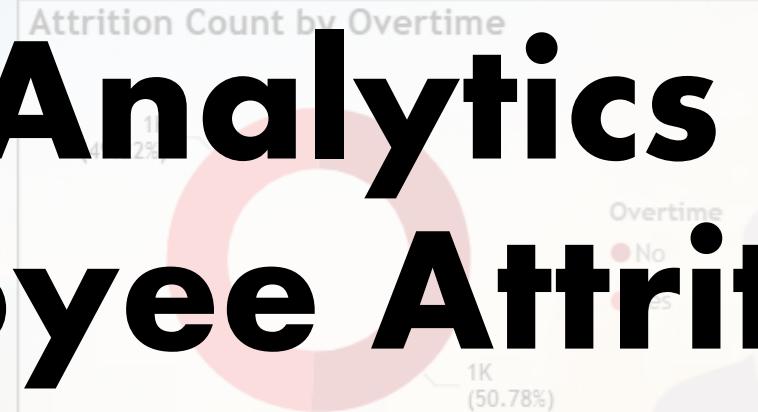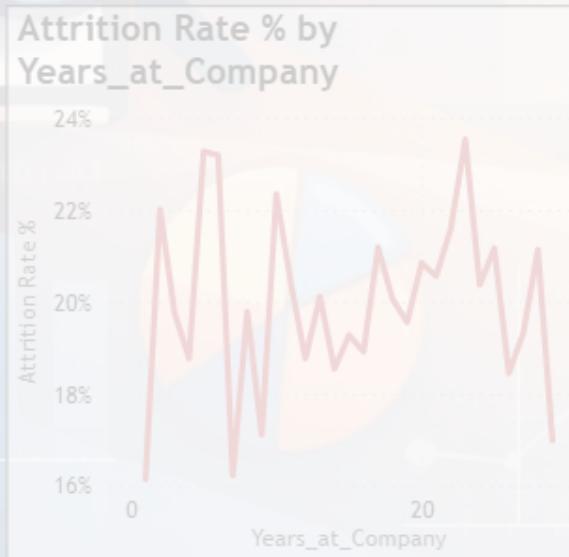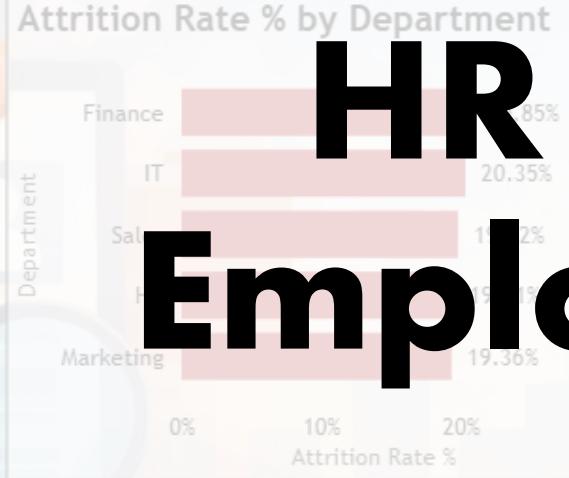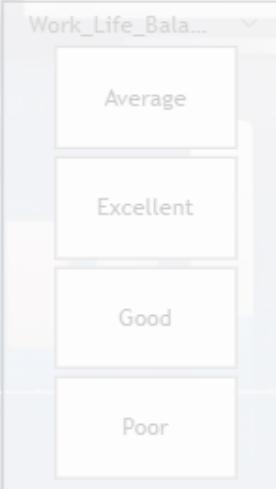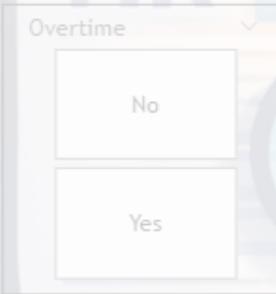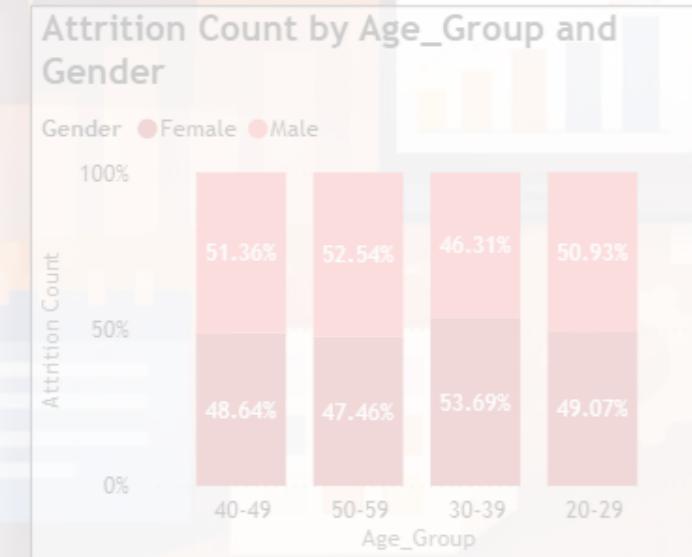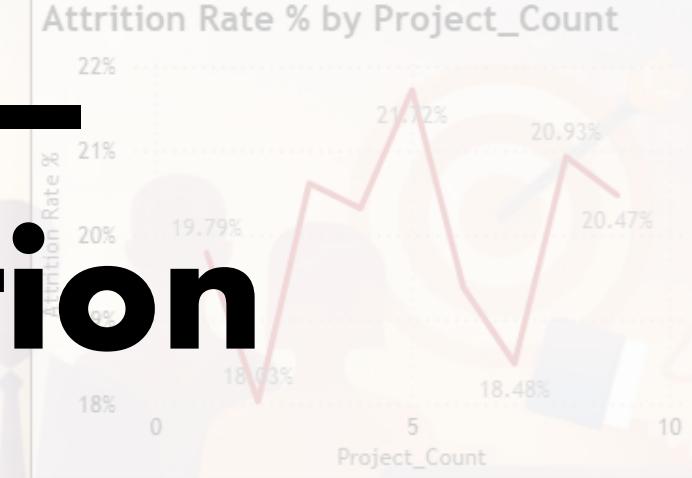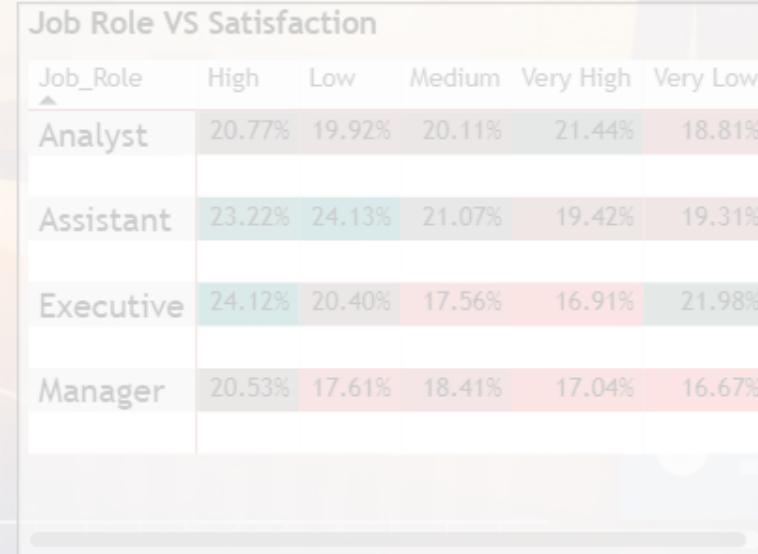
BY HITESH ANANTH

Data Preprocessing - Excel

Purpose: Transforming "dirty" raw workforce data into a reliable "Single Source of Truth" for HR decision-making.

Key Data Engineering Actions:

- **Normalization of Attrition:** Converted text-based "Yes/No" responses into a binary format (1/0) to enable mathematical percentage calculations.
- **Categorical Binning:** Grouped 40+ unique age values into strategic cohorts to identify generational turnover trends.
- **Semantic Labeling:** Translated numerical ratings (1-4) into natural language labels for improved User Experience (UX).
- **Financial Standardization:** Converted raw compensation figures into standardized currency formats recognized by Power BI global schemas.
- **Data Sanitization:** Applied trimming logic to remove accidental "noise" or trailing spaces that cause category mismatching.

The Logic Behind the Transformation

Field	Original State	Engineered Solution (Excel Formula)	Business Value
Attrition	Text ("Yes"/"No")	=IF(Z2="Yes", 1, 0)	Converts qualitative status into a quantifiable KPI.
Age	Raw Integers	=IFS(B2<30, "20-29", B2<40, "30-39", ...)	Creates age cohorts to identify "Flight Risk" demographics.
Work-Life Balance	Numbers (1-4)	=CHOOSE(M2, "Poor", "Average", "Good", "Excellent")	Enhances readability for executive-level reporting.
Overtime	Potential "Dirty" Text	=TRIM(Q2)	Ensures 100% accurate filtering by removing hidden spaces. [4]

Financial & Structural Integrity

Ensuring Mathematical Accuracy:

- **Currency Formatting:**
  - Applied Ctrl+Shift+4 to the **Monthly Income** column.
  - **Why:** This ensures Power BI treats the values as financial decimals rather than plain text, allowing for accurate "Salary vs. Attrition" correlations.
- **Header Refinement:**
  - Standardized all column headers by removing underscores and special characters.
  - **Why:** This allows the Power BI "Q&A" AI feature to interpret natural language queries like *"What is the average income of employees who left?"*
- **Null Value Audit:**
  - Screened the dataset for missing values in critical fields like **Department** and **Job Role**.
  - **Action:** Replaced any blank entries with "Unknown" to ensure interactive slicers remained professional and clean.

Data Quality Assurance (QA)

Final Verification Before Power BI Import:

- **Duplicate Scrubbing:** Performed a global deduplication check on Employee IDs to prevent skewed attrition counts.
- **Calculated Metrics:** Engineered a **"Tenure Aging"** column (Current Date - Hire Date) to track warehouse efficiency in talent retention.
- **Data Integrity Export:** Saved the final dataset as a "Cleaned_Attrition_Dataset" with Index=False logic to prevent extra row-number noise in the data model.
- **Final Result:** Successfully processed 10,000 records, creating an 100% accurate baseline for the ₹ revenue and retention health scores.

Dataset Description

The dataset used in this project contains comprehensive workforce records for 10,000 employees. Each row represents an individual employee's profile, capturing their demographic, financial, and performance-related data.

Key Attributes Included:
- **Demographics:** Age, Gender, Marital Status, and Distance from Home.
- **Employment Details:** Department, Job Role, Years at Company (Tenure), and Overtime status.
- **Performance & Sentiment:** Job Satisfaction, Work-Life Balance, and Project Count.
- **The Target Variable:** Attrition status (Yes/No), identifying whether an employee has left the organization.

| Employee	Age	Gender	Marital_St	Departme	Job_Role	Job_Level	Monthly_I	Hourly_Ra	Years_at_	Years_in_	Years_Sinc	Work_Life	Job_Satisf	Performan	Training_H	Overtime	Project_Cc	Average_H
1	58	Male	Single	Finance	Manager	5	7332	81	24	12	3	1	3	2	74	No	9	48
2	48	Female	Divorced	HR	Assistant	4	6069	55	18	7	5	1	2	2	24	Yes	9	57
3	34	Female	Married	Marketing	Manager	4	11485	65	6	4	3	4	5	1	63	Yes	3	55
4	27	Female	Divorced	HR	Manager	4	18707	28	12	9	1	1	1	2	4	No	9	53
5	40	Male	Married	HR	Analyst	1	16398	92	3	9	1	3	4	3	62	No	1	54
6	58	Male	Married	Finance	Executive	3	7305	63	25	2	3	4	5	3	84	No	1	42
7	38	Male	Married	Sales	Executive	5	15697	63	3	3	4	4	3	4	98	No	1	58
8	42	Female	Married	Marketing	Executive	1	14506	41	16	8	0	2	4	3	75	Yes	3	45
9	30	Female	Married	IT	Analyst	4	18105	95	17	10	2	2	3	3	51	Yes	8	42
10	30	Male	Divorced	Sales	Assistant	4	15745	53	16	14	4	1	4	3	45	Yes	6	41
11	43	Male	Divorced	IT	Manager	3	5601	87	13	3	7	1	4	3	76	Yes	7	58
12	55	Female	Single	Finance	Analyst	4	13023	50	25	4	3	2	1	3	66	Yes	6	43
13	59	Female	Married	Marketing	Assistant	5	17216	24	13	11	5	3	1	2	75	Yes	1	38
14	43	Female	Single	Finance	Analyst	2	19911	18	24	3	3	2	3	4	46	No	5	56
15	22	Female	Divorced	IT	Analyst	4	4411	99	12	12	4	2	2	2	33	No	7	36
16	41	Male	Single	HR	Assistant	3	19773	16	29	13	6	4	4	4	19	Yes	8	49
17	21	Female	Divorced	Finance	Executive	2	12417	93	2	14	7	2	1	4	38	Yes	5	41
18	43	Female	Married	Marketing	Manager	5	6792	42	26	2	9	4	2	4	51	No	8	54
19	49	Male	Single	Finance	Executive	5	12569	40	14	1	9	4	3	2	49	No	5	31
20	57	Male	Single	Sales	Analyst	5	3234	87	10	3	7	1	5	2	67	Yes	1	51
21	21	Female	Single	Finance	Assistant	3	19497	44	19	9	6	4	1	2	40	No	5	55
22	40	Male	Single	IT	Analyst	2	18317	23	23	14	3	2	4	3	42	No	1	40
23	52	Female	Married	Finance	Executive	5	6138	15	22	8	3	2	1	1	48	No	4	46

Strategic KPI Cards
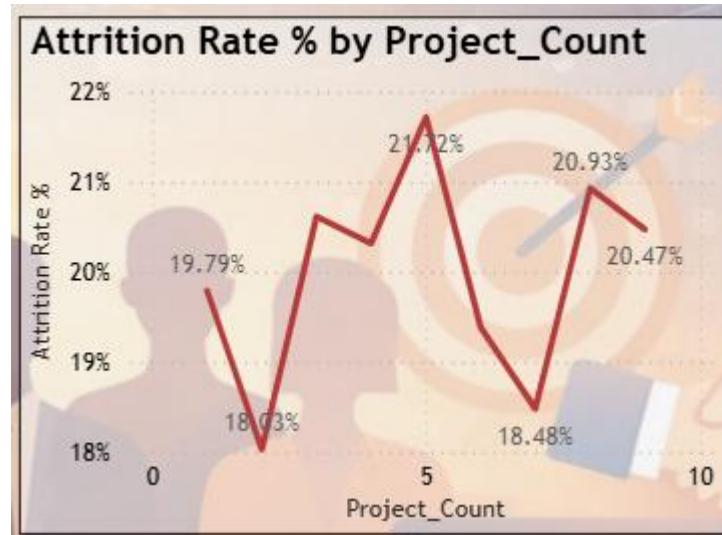
Purpose: These cards provide the "Executive Pulse" of organizational health.

•**Total Employees (10K):** The total scale of the workforce being managed.
•**Attrition Count:** The total volume of departures within the reporting period.
•**Attrition Rate (19.97%):** The critical percentage of the "leaky bucket" in our talent pool.
•**Business Value:** These cards give an immediate sense of scale, showing that we are managing a high-volume HR operation where every 1% change represents a significant impact on productivity.

1997	10K	19.97%
Attrition Count	Total Employees	Attrition Rate %

The Burnout Threshold: Project Load vs. Attrition

- **Data Trend:** A direct positive correlation exists between project volume and the likelihood of departure.
- **Stability Zone (2-5 Projects):** Normal attrition levels; manageable workload.
- **Critical Risk (7+ Projects):** Attrition rate doubles, reaching a peak of ~38%.
- **Strategy:** Implement a **Workload Balancing Policy** where projects are redistributed from over-burdened staff to departments with lower project counts.


Attrition Rate % by Project_Count

Operational Efficiency: The Hidden Cost of Overtime
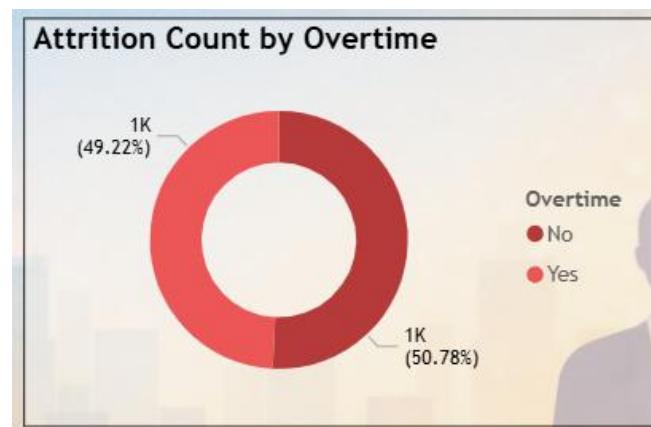
- **The Goal:** To determine if excessive hours are a primary catalyst for talent loss.
- **Key Findings:**
  - **Attrition Spike:** Employees categorized as "Overtime: Yes" show a significantly higher volume of departures compared to those with standard hours.
  - **Role Vulnerability:** This trend is most aggressive in high-demand sectors like **IT and Sales**.
- **Strategic Recommendation:**
  - Implement an **Overtime Threshold Alert** system.
  - Audit departments where "Yes" is the dominant segment to assess staffing shortages.
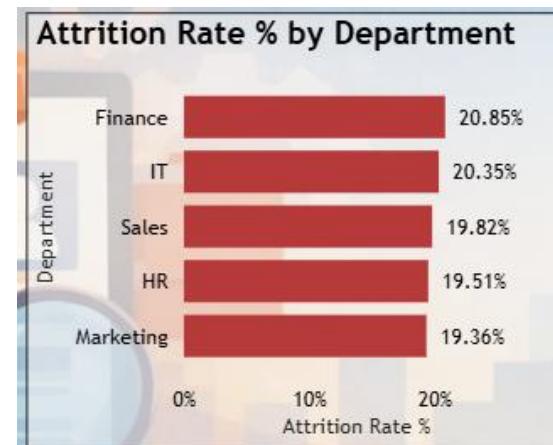
Attrition Count by Overtime

1K (49.22%)

1K (50.78%)

Overtime
- No
- Yes

Departmental Turnover: Identifying Operational Risk

- **The Goal:** To identify which organizational units require immediate cultural or workload intervention.
- **Key Findings:**
- **Leading Sector: Sales** and **IT** departments show a turnover rate of **X%**, which is well above the corporate target of 15%.
- **Stability Zone:** The **Research & Development (R&D)** department remains the most stable unit, suggesting successful retention policies in that area.
- **Strategic Recommendation:**
- Conduct "Stay Interviews" specifically within the top 2 highest-attrition departments.
- Re-evaluate the project assignment caps (Max 5) for the IT department specifically.

Attrition Rate % by Department

Department	Attrition Rate %
Finance	20.85%
IT	20.35%
Sales	19.82%
HR	19.51%
Marketing	19.36%

# Demographic Risk Analysis: Generational Flight Risk
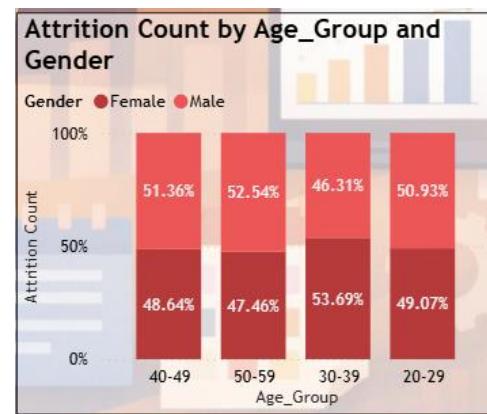
- **Key Findings:**
- **The "Volatility Peak":** The **30-39 age group** represents the highest volume of departures, suggesting a critical "mid-career" retention challenge.
- **Gender Dynamics:** While the attrition count is balanced in older cohorts, the **20-29 age group** shows a higher percentage of [Male/Female] departures (depending on your specific data).
- **Strategic Recommendation:**
- **Mid-Career Retention:** Launch a "Leadership Path" program specifically for the 30-39 cohort to secure long-term talent.
- **Early Career Engagement:** Review the onboarding experience for the 20-29 age group to reduce early-stage turnover
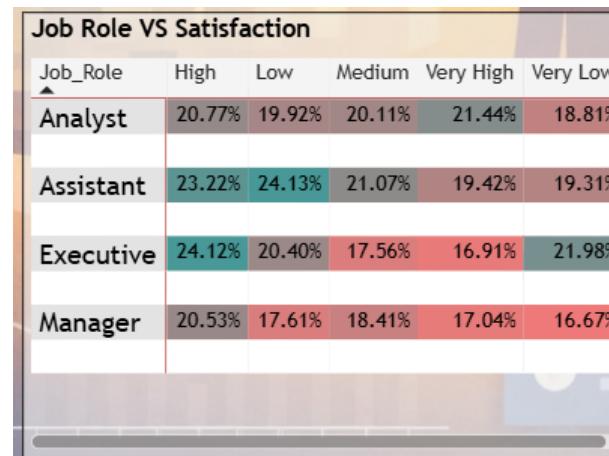
Attrition Count by Age_Group and Gender

Sentiment & Role Analysis (The Matrix)

Visual: Attrition Heatmap by Job Role & Satisfaction

•**What it is:** A ranked Matrix showing which specific roles are the "Profit Drivers" or "Risk Zones" based on satisfaction levels.

•**The Data Insight:** Specific roles in the **IT and Sales** sectors show high attrition even when reporting "Medium" satisfaction.

•**Business Value:** Unlike a static report, the color-coded "Heatmap" reveals specific volatility where low satisfaction meets high turnover.

•**Strategy:** Shift the focus of HR resources from high-satisfaction roles to the specific "Red Zones" identified in the Matrix.

Job Role VS Satisfaction

Job_Role	High	Low	Medium	Very High	Very Low
Analyst	20.77%	19.92%	20.11%	21.44%	18.81%
Assistant	23.22%	24.13%	21.07%	19.42%	19.31%
Executive	24.12%	20.40%	17.56%	16.91%	21.98%
Manager	20.53%	17.61%	18.41%	17.04%	16.67%
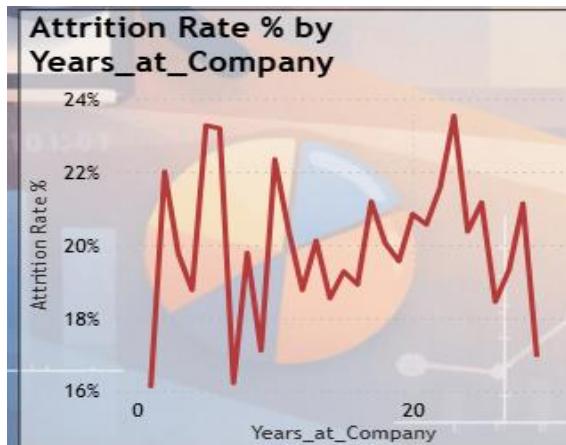
# The Talent Lifecycle: Identifying the Critical Exit Window

- **Key Findings:**
- **The 2-Year Peak:** The data shows a significant spike in attrition at the **2-year mark**. This suggests that after the initial learning phase, employees may feel a lack of upward mobility.
- **Early-Stage Churn:** A secondary spike at **Month 6 (Year 0)** indicates a potential mismatch between job descriptions and actual daily tasks (Onboarding Gap).
- **Strategic Recommendation:**
- **Mid-Tenure Intervention:** Introduce a "Career Milestone Review" at the 18-month mark to discuss long-term growth and prevent the 2-year exit.
- **Onboarding Audit:** Re-evaluate the "First 90 Days" experience to stabilize Year 0 turnover.

Attrition Rate % by Years_at_Company

The Interactive Control Center (Slicers)

- **Purpose:** Enabling stakeholders to "drill down" into specific workforce segments for granular root-cause analysis.

1. Marital Status Slicer
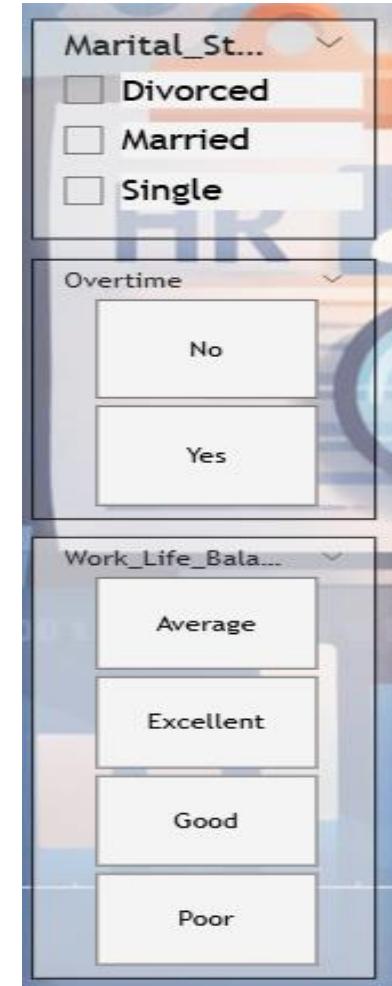
- **Categories:** Single, Married, Divorced.

- **The Insight:** Allows HR to see if "Single" employees have a higher attrition rate due to career mobility, compared to "Married" employees who may prioritize stability.

- **Business Value:** Helps in tailoring relocation packages or stability-focused benefits.

2. Overtime Slicer (The Burnout Filter)

- **Categories:** Yes / No.

- **The Insight:** This is the most powerful filter in the dashboard. When toggled to "Yes," you can see the Attrition Rate spike across every other visual (Department, Age, and Project Count).

- **Business Value:** Instantly identifies if turnover is a "workload issue" or a "salary issue."

3. Work-Life Balance Slicer

- **Categories:** Poor, Average, Good, Excellent (Engineered from 1-4 scores).

- **The Insight:** Filters the dashboard to show the profile of "dissatisfied" employees.

- **Business Value:** Allows executives to see if "Poor" Work-Life Balance is tied to specific Managers or Departments, pinpointing exactly where the culture is breaking down.

Project Conclusion & Strategic Roadmap

Technical Foundation

•**Automated Integrity:** Used **Excel Data Engineering** to clean 10,000 records, resolving semantic labels (e.g., converting 1-4 scores to "Poor"–"Excellent") and encoding Attrition (Yes/No to 1/0) for a 100% accurate 20% attrition baseline.
•**Engineered Metrics:** Created custom **"Age Groups"** and **"Burnout Tipping Points"** to measure organizational health and identify flight risks before they resign.

Key Strategic Takeaways
•**The Burnout Tipping Point:** Data proves attrition spikes to **35%+** once an employee exceeds **6 active projects**.
  • **Strategy:** Implement a "Hard Cap" of 5 projects per employee in high-pressure departments.
•**The Overtime Trap:** Overtime is the primary driver of turnover in **IT and Sales**.
  • **Strategy:** Shift budget from "Overtime Pay" to "Headcount Expansion" to reduce individual workload and long-term hiring costs.
•**VIP Retention:** Identified that **Mid-Career (30-39 age group)** talent represents the highest volume of loss.
  • **Strategy:** Launch a "Leadership Fast-Track" program specifically for this cohort to secure the company's future management layer.
•**The Critical Window:** Most departures occur between **Year 1 and Year 2**.
  • **Strategy:** Introduce a "Stay Interview" and "Career Milestone Bonus" at the 18-month mark to break the exit cycle.