# Report of Project

# Phase II: Decision Making

By [Asha Mishra](mailto:amish140@asu.edu), [Taljinder Singh](mailto:tsingh80@asu.edu), [Jacob Kuriakose](mailto:jkuriak3@asu.edu), [Hitha Mahesh Gowda](mailto:hgowda4@asu.edu)

Contents:

* Section 1: Used Visualization Tools
* Section 2: Explanation of Required Data Pre-processing, if any
* Section 3: List of Final Sets of Questions
* Section 4: Dashboard Plot Drafts
* Section 5: Dashboard Interactivity
* Section 6: References, Include link to team's Mural board

Used Visualization Tools

**Tableau Prep Builder**

* **Why:** Native support for cleaning, type‐casting, binning, pivoting, and dropping fields via a visual flow. Exports Hyper/CSV extracts that feed directly into Tableau Desktop.

**Tableau Desktop**

* **Why:** Powerful drag-and-drop interface for creating interactive dashboards, dual-axis and small-multiples layouts, and built-in interactivity (filters, parameters). Ideal for rapid prototyping and stakeholder demos.

**TabPy (Python-in-Prep)**

* **Why:** Only used for highly custom calculations (e.g. complex imputation). All standard prep steps are handled natively, avoiding extra scripting overhead.

Explanation of Required Data Pre-processing, if any

All prep was performed in Tableau Prep Builder on the original HR Analytics CSV, producing Output.csv. Below are the details of the pre-processing steps performed:

**Dropped non-informative or constant fields:** EmpID, EmployeeNumber, EmployeeCount, MonthlyRate, Over18, SalarySlab, StandardHours

**Attrition:** Mapped Yes/No → 1/0

**BusinessTravel:** Clean labels; changed Travelrarely to Travel\_rarely

**Education:** Convert to categorical (1 = Very Less Educated → 5 = Very High Educated).

**EnvironmentSatisfaction:** Convert to categorical (1 = Low, 2=Medium, 3=High, 4 = Very High).

**JobInvolvement:** Convert to categorical (1 = Low, 2=Medium, 3=High, 4 = Very High).

**JobLevel:** Convert to categorical (1 - Very Low, 5- Very High).

**JobSatisfaction:** Convert to categorical (1 = Low, 2=Medium, 3=High, 4 = Very High).

**OverTime:** Mapped Yes/No → 1/0

**RelationshipSatisfaction:** Convert to categorical (1 = Low, 2=Medium, 3=High, 4 = Very High).

**WorkLifeBalance:** Convert to categorical (1 = Low, 2=Medium, 3=High, 4 = Very High).

**YearsWithCurrManager:** Fill missing values with median

List of Final Sets of Questions

**Question 1:** What is our overall attrition rate, and how does it differ by Department and JobRole?

**Question 2:** How do EducationFields and Education levels relate to PerformanceRating?

**Question 3:** What patterns emerge between PercentSalaryHike and subsequent PerformanceRating?

**Question 4:** How do attrition rates vary across AgeGroup and Gender?

**Question 5:** How do the four Satisfaction scores (Environment, Job, Relationship, Work-Life) compare by Department and how do they predict Attrition?

**Question 6:** Which JobLevels and JobRoles show the strongest PerformanceRating and lowest Attrition?

**Question 7:** Is lower JobInvolvement associated with higher Attrition?

**Question 8:** Is there a relationship between MonthlyIncome (or PercentSalaryHike) and Attrition?

**Question 9:** Does TrainingTimesLastYear drive higher PerformanceRating?

**Question 10:** How do BusinessTravel categories (Rarely, Frequently, None) affect WorkLifeBalance and Attrition?

**Question 11:** How does DistanceFromHome correlate with OverTime and Attrition?

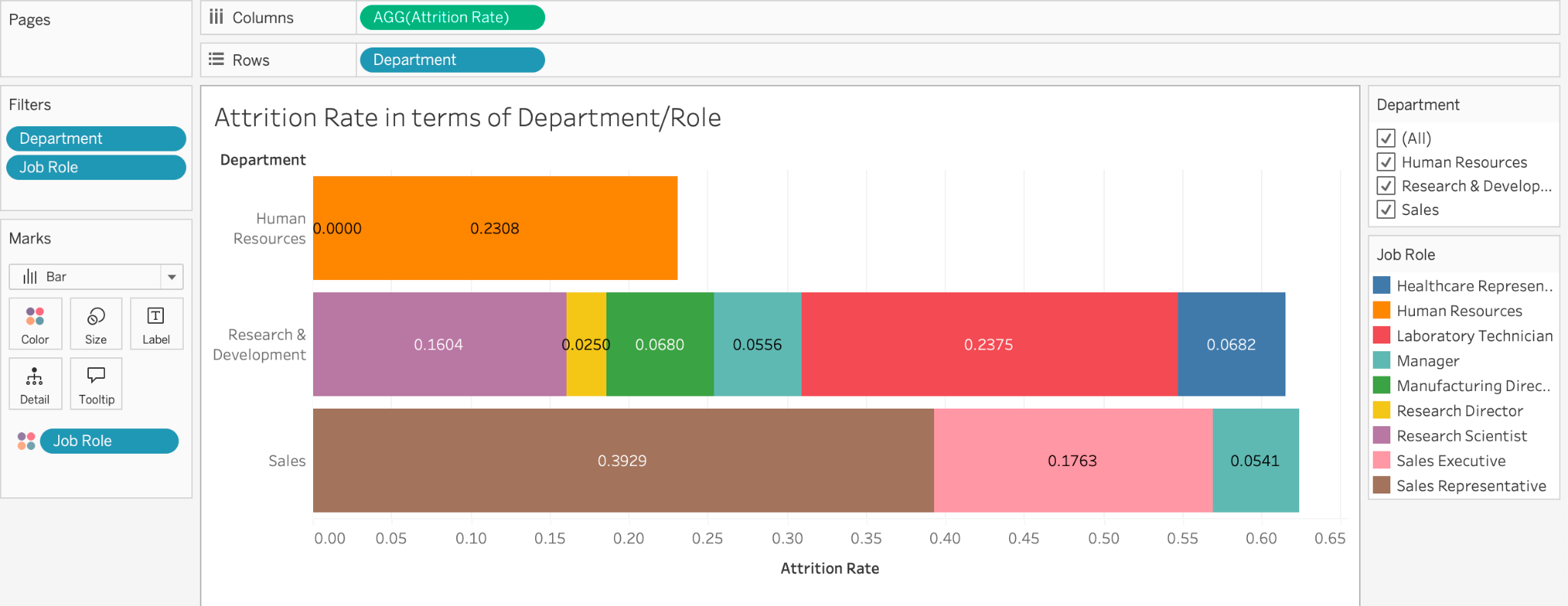
**Question 12:** How does YearsSinceLastPromotion impact both Attrition risk and PerformanceRating?

**Question 13:** What’s the distribution of StockOptionLevel, and do richer equity packages boost retention?

**Question 14:** Are there gender‐based differences in pay rates (HourlyRate, DailyRate, MonthlyIncome) that correlate with Attrition?

Dashboard Plot Drafts

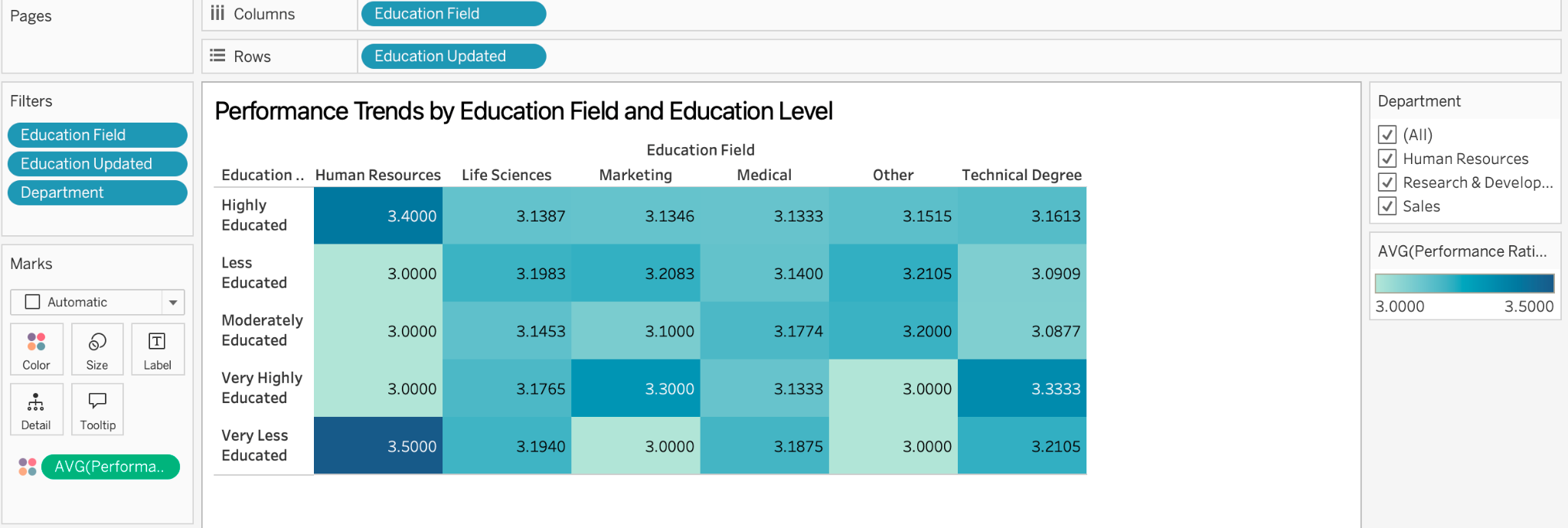
**Question: 1** What is our overall attrition rate, and how does it differ by Department and JobRole?

**  
Pre-attentive attributes:** Color: Job Role

**What it shows:** Attrition Rate in terms of Department and Job Role

**How it relates:** A stacked bar is used to show varied attrition rates from different job roles in various departments.

**Question 2:** How do EducationFields and Education levels relate to PerformanceRating?



**Pre-attentive attributes:** Color: Performance Rating

**What it shows:** It shows different education backgrounds vs their incline/lean of their performance rating.

**How it relates:** A heat map of squares

**Question 3: What patterns emerge between PercentSalaryHike and subsequent PerformanceRating?**

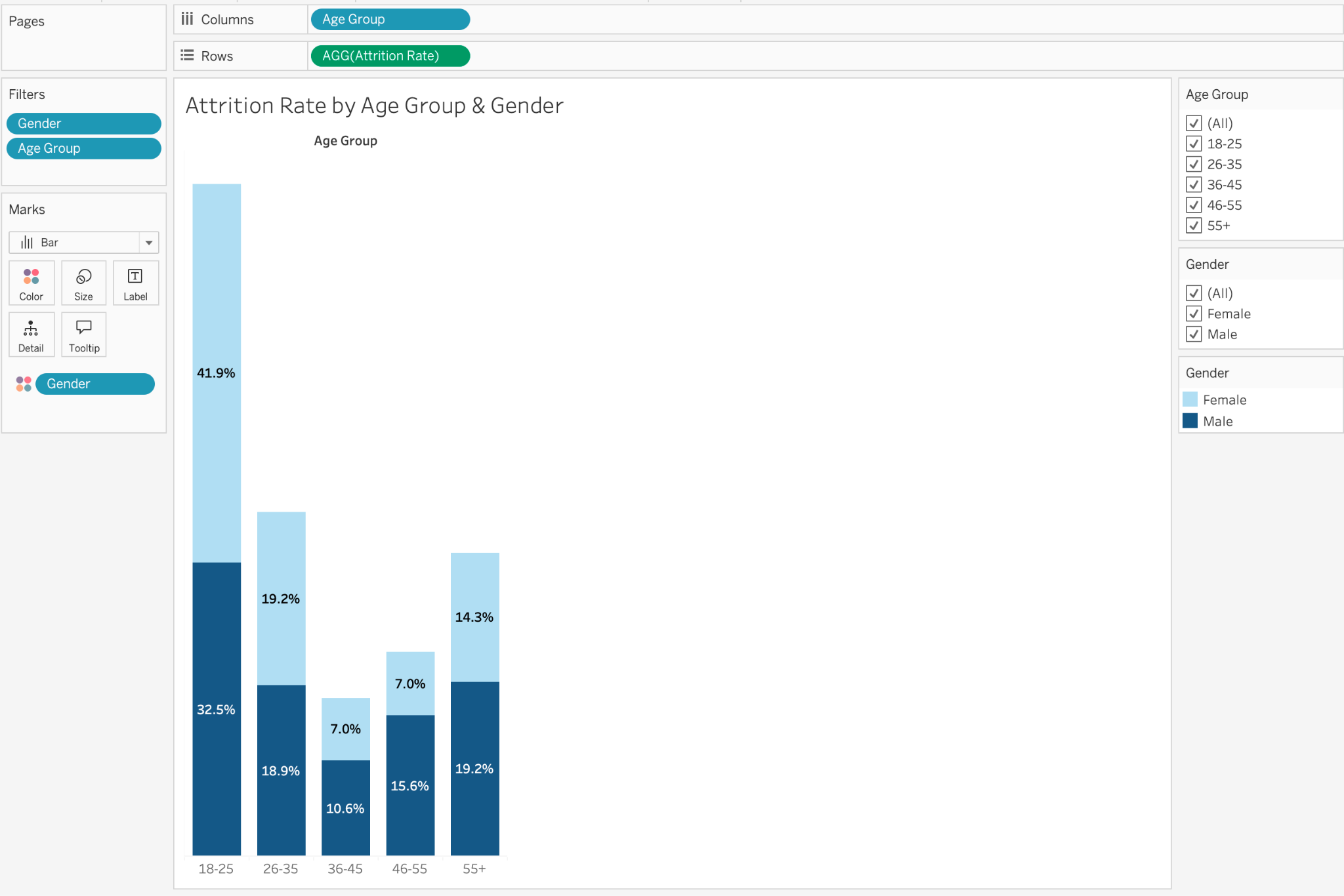


**Preattentive Attributes:** Color:Job Role

**What it shows:** This line graph plots Performance Rating against Percent Salary Hike across different Job Roles.

**How it relates:** A clear pattern is that moderate salary hikes (~12%-14%) correlate with higher performance, but beyond that, increased salary hikes don't necessarily mean higher performance and might even decrease it.

**Question 4: How do attrition rates vary across AgeGroup and Gender?**

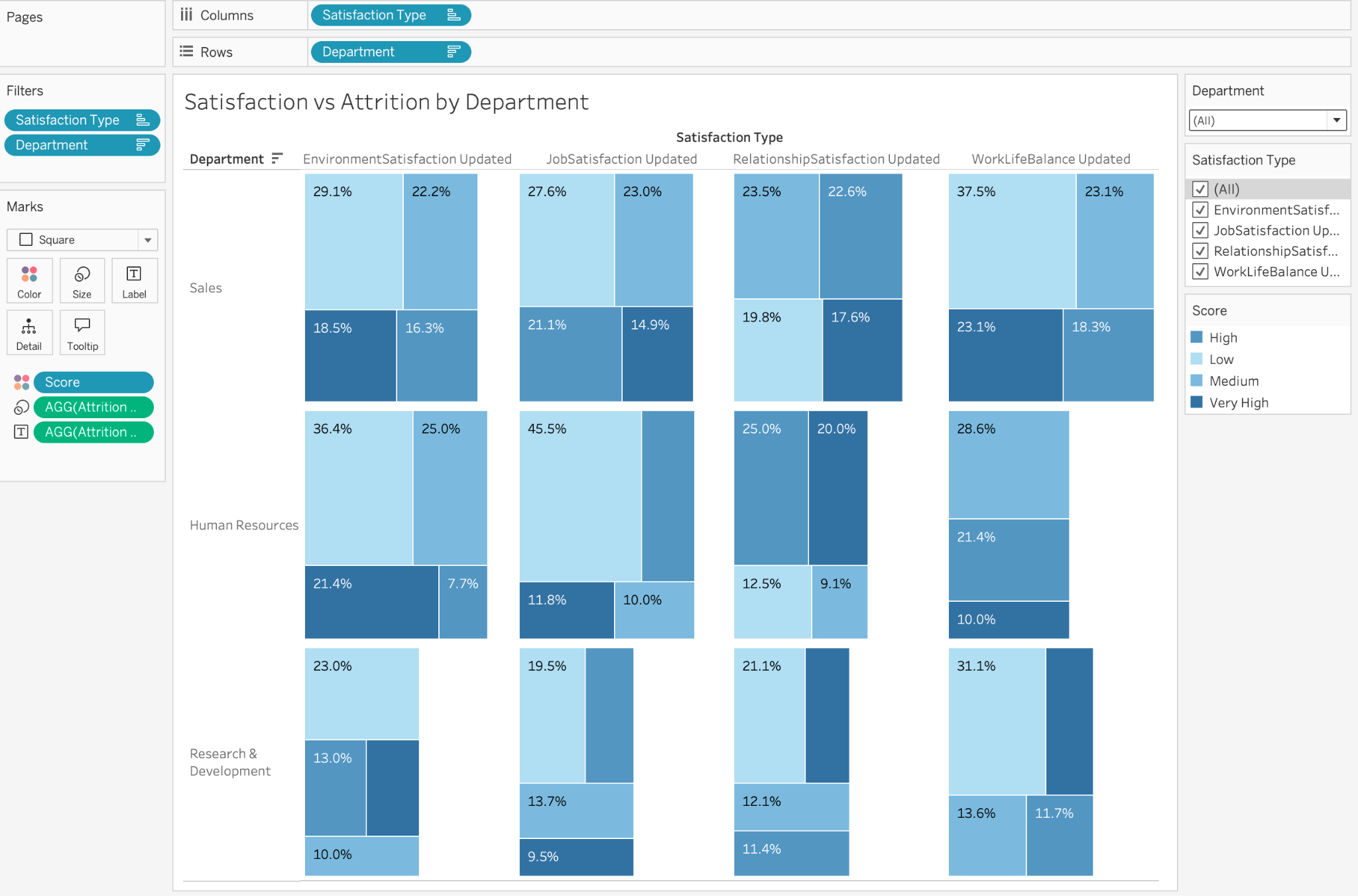


**Pre-attentive attributes:** Bar Length = Attrition Rate; Color = Gender

**What it shows:** A stacked‐bar chart with one bar per AgeGroup (18–25, 26–35, …, 55+). Each bar’s total height is the overall attrition rate for that cohort, and the bar is stacked into two segments, dark blue for male attrition and light blue for female attrition. You can quickly compare both the total turnover in each age bucket and the gender split within it.

**How it relates:** This view exemplifies our demographic segmentation of attrition. It lays the groundwork for all other attrition analyses by showing how basic personal attributes (age, gender) correlate with leaving, before we dive into job, salary, engagement or performance factors.

**Question 5: How do the four Satisfaction scores (Environment, Job, Relationship, Work-Life) compare by Department and how do they predict Attrition?**

****

**Pre-attentive attributes:** Position = Dept × SatisfactionType grid; Color = average satisfaction score (light→dark); Size = attrition rate (small→large)

**What it shows:** A heatmap of squares in a grid with Departments on the y-axis and the four satisfaction types on the x-axis. Each cell’s color encodes average satisfaction (light→dark) and size encodes that department’s attrition rate. You instantly spot, for each survey metric where satisfaction is low and attrition is high.

**How it relates:** This combines our engagement metrics (the four satisfaction scores) with attrition, while also cross‐linking to other “drivers” visualizations (training, promotions). It’s our first multi-metric view, showing how subjective well-being maps onto actual turnover.

**Question 6: Which JobLevels and JobRoles show the strongest PerformanceRating and lowest Attrition?**

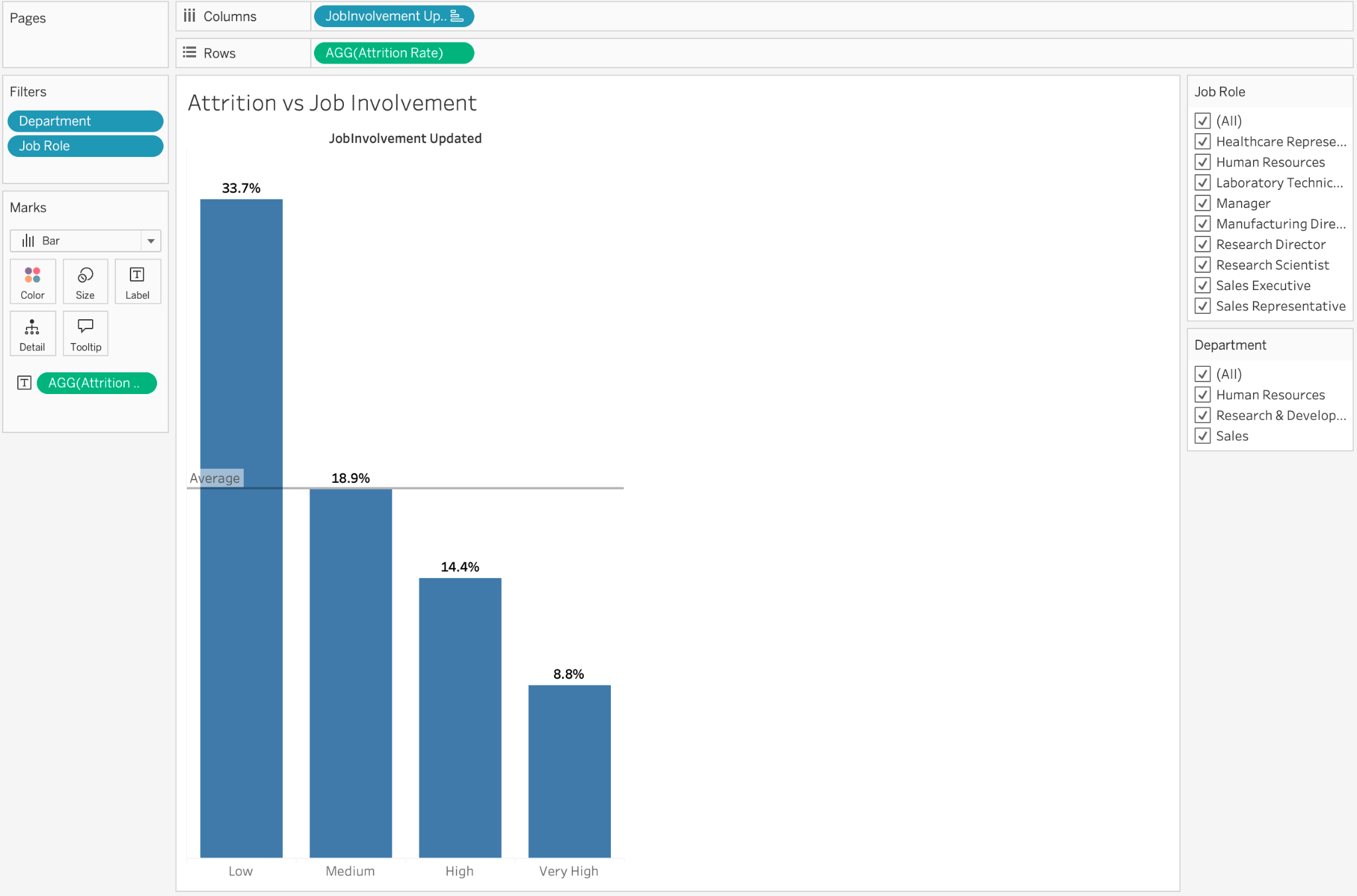


**Pre-attentive attributes:** Position = JobLevel & JobRole grid; Size = Attrition Rate (small→large); Color = Avg Performance

**What it shows:** A bubble-matrix: a grid where the x-axis is Job Level and the y-axis is Job Role. Each circle’s size is that cell’s attrition rate and its color is the average performance rating (cold→warm). You can quickly identify the “sweet spots” (small bubbles) and risk areas (large bubbles).

**How it relates:** This is our performance vs. retention cross-analysis (complementing education, salary hike, and experience visualizations). It shows which combinations of seniority and role yield both high performance and strong retention, guiding promotion and staffing decisions.

**Question 7: Is lower JobInvolvement associated with higher Attrition?**

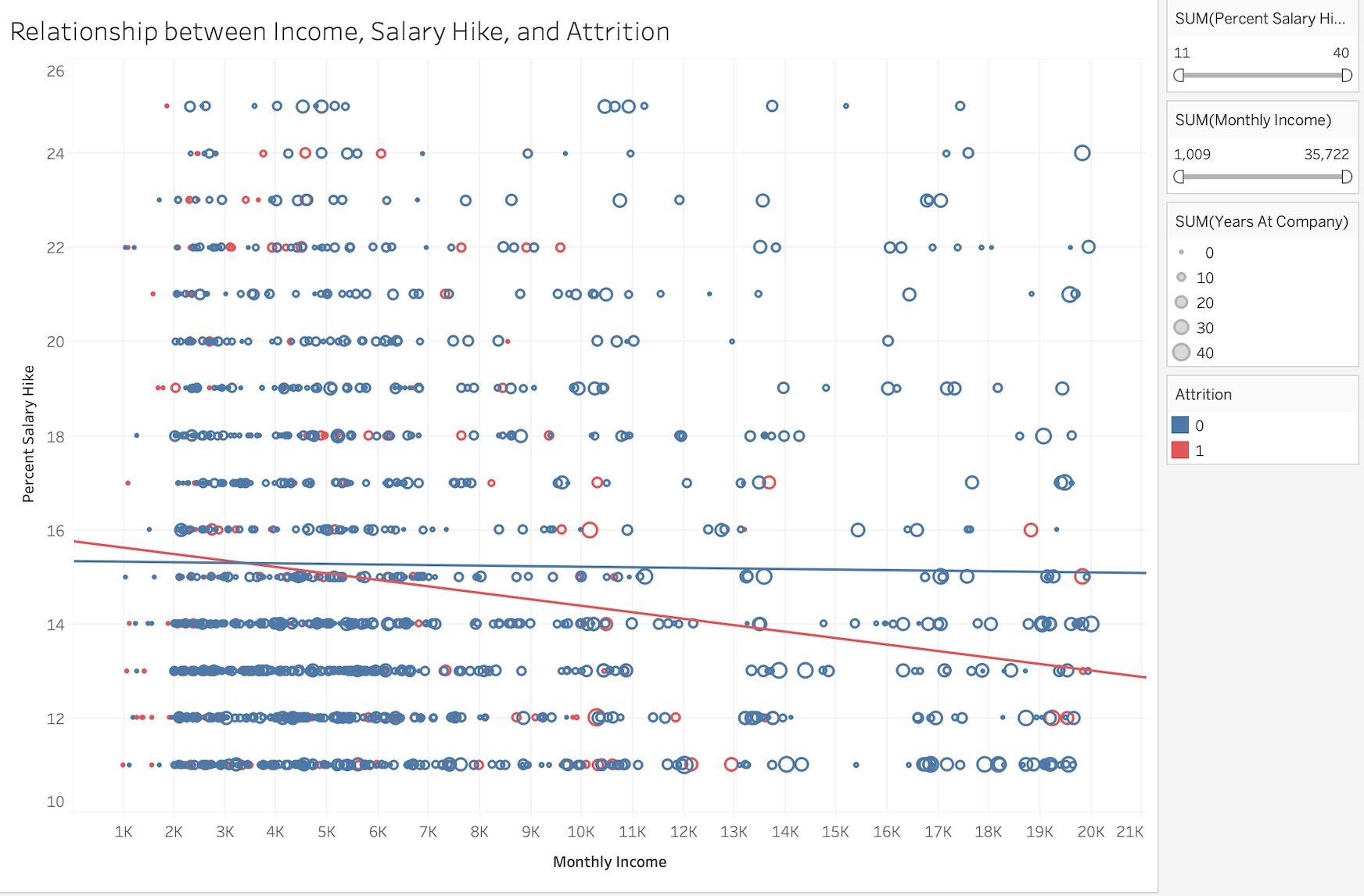
****

**Pre-attentive attributes:** Bar length = Attrition Rate

**What it shows:** A simple bar chart with JobInvolvement levels on the x-axis (Low → Medium → High → Very High) and bar heights equal to each level’s attrition rate. As the involvement is truly protective, you see the tallest bars on the left (lowest involvement).

**How it relates:** This addresses our engagement/behavior cluster, alongside satisfaction and training. It isolates one engagement metric (JobInvolvement) to confirm that employees who feel less connected to their work indeed leave at higher rates, reinforcing the broader picture of what drives attrition.

**Question 8: Is there a relationship between MonthlyIncome (or PercentSalaryHike) and Attrition?**



### **Pre-attentive attributes:**

Color = Attrition;  
Size = Years at Company;  
Position = Monthly Income vs Percent Salary Hike.

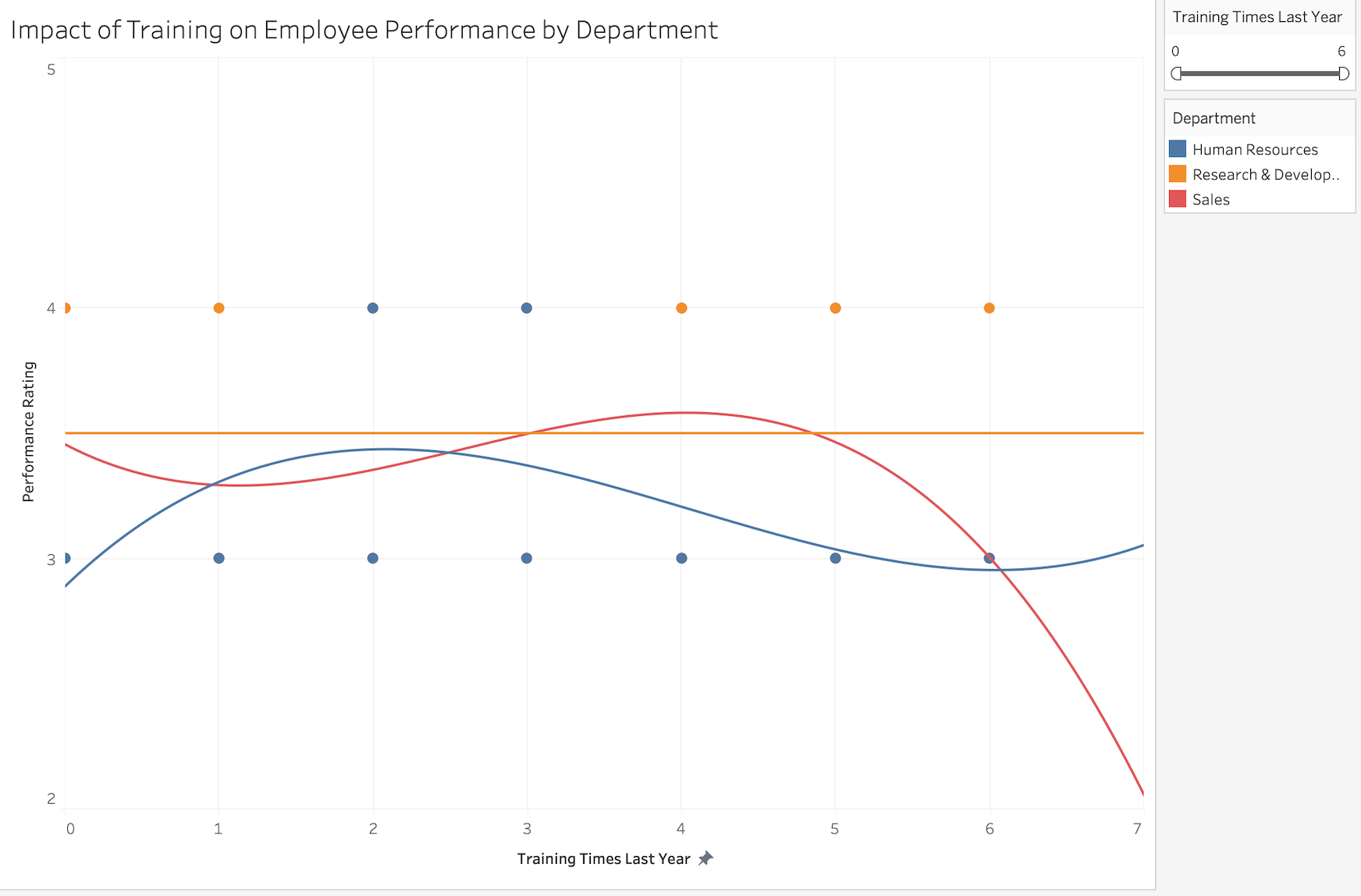
### **What it shows:**

A scatterplot with Monthly Income on the x-axis and Percent Salary Hike on the y-axis.  
Each point's color indicates Attrition (red = left, blue = stayed), and the size of the point shows the number of years at the company.  
The added polynomial trend line helps detect if there’s any consistent relationship between these financial metrics and employee attrition.

### **How it relates:**

This visualization answers the question about whether financial rewards (monthly income or salary hikes) correlate with employees leaving the company.  
By seeing mostly flat or weak trends, it suggests that income or salary hikes alone are **not major predictors** of attrition — other factors like engagement or work-life balance might be stronger drivers.  
It helps narrow the focus away from financial compensation as the primary reason for turnover.

**Question 9: Does TrainingTimesLastYear drive higher PerformanceRating?**



### **Pre-attentive attributes:**

Color = Department;  
 Position = Training Times Last Year vs Performance Rating;  
 Curve Shape = Polynomial fitted trend line per department.

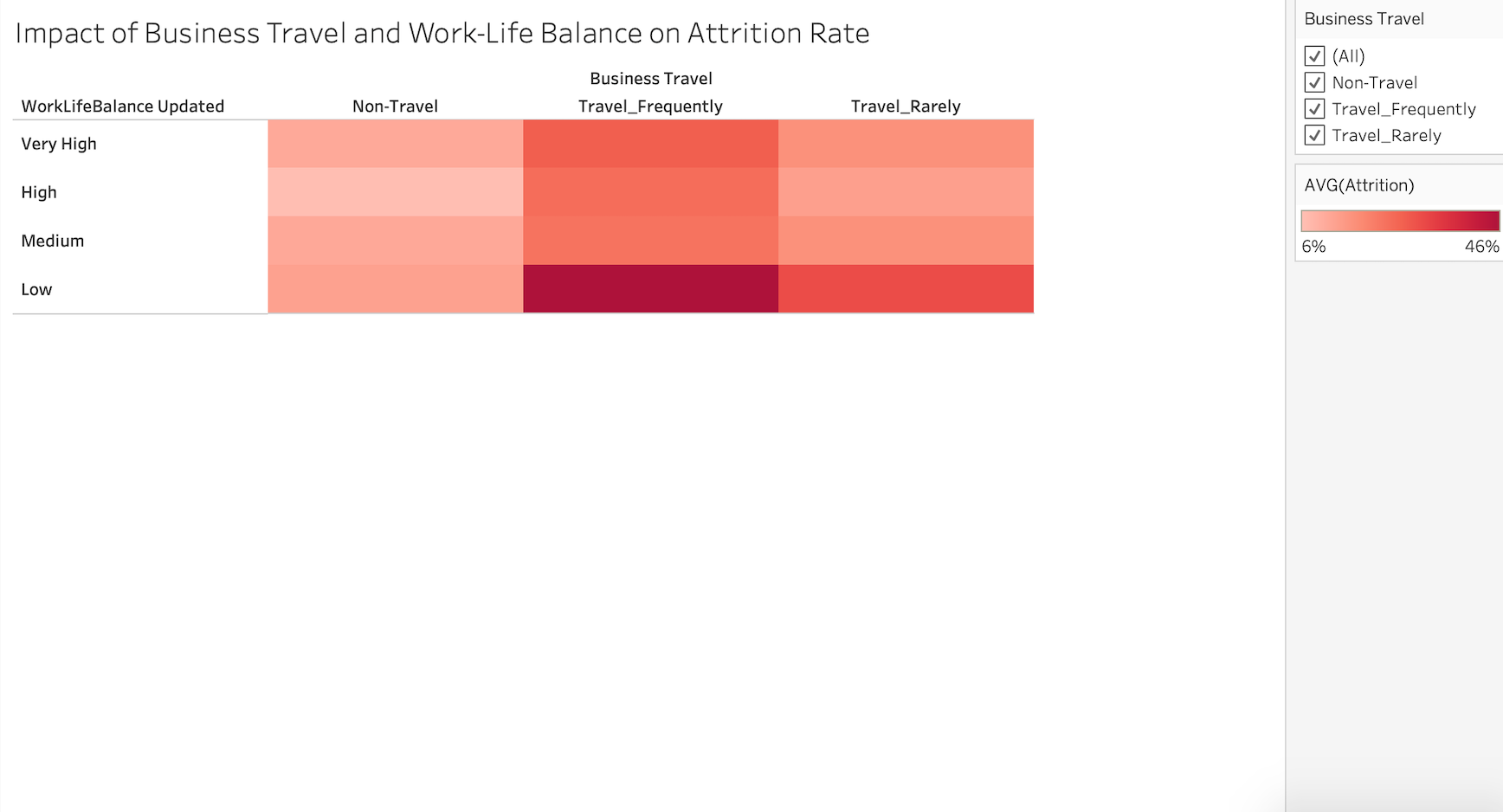
### **What it shows:**

A scatterplot plotting the number of training sessions last year against the employee’s performance rating, separated by Department.  
 Each department has its own polynomial trend line showing how training relates to performance.  
 For instance, Human Resources has a weaker trend compared to Sales or Research & Development.

### **How it relates:**

This visualization addresses whether training frequency is **positively correlated with better performance ratings** — a key hypothesis about employee development and outcomes.  
 The differing slopes between departments show that **training might benefit some departments more than others** (like Sales and R&D seeing gains), guiding where training investments might yield better performance improvements.

**Question 10: How do BusinessTravel categories (Rarely, Frequently, None) affect WorkLifeBalance and Attrition?**



### **Pre-attentive attributes:**

Color Intensity = Attrition Rate;  
 Grid Position = Business Travel vs Work-Life Balance levels.

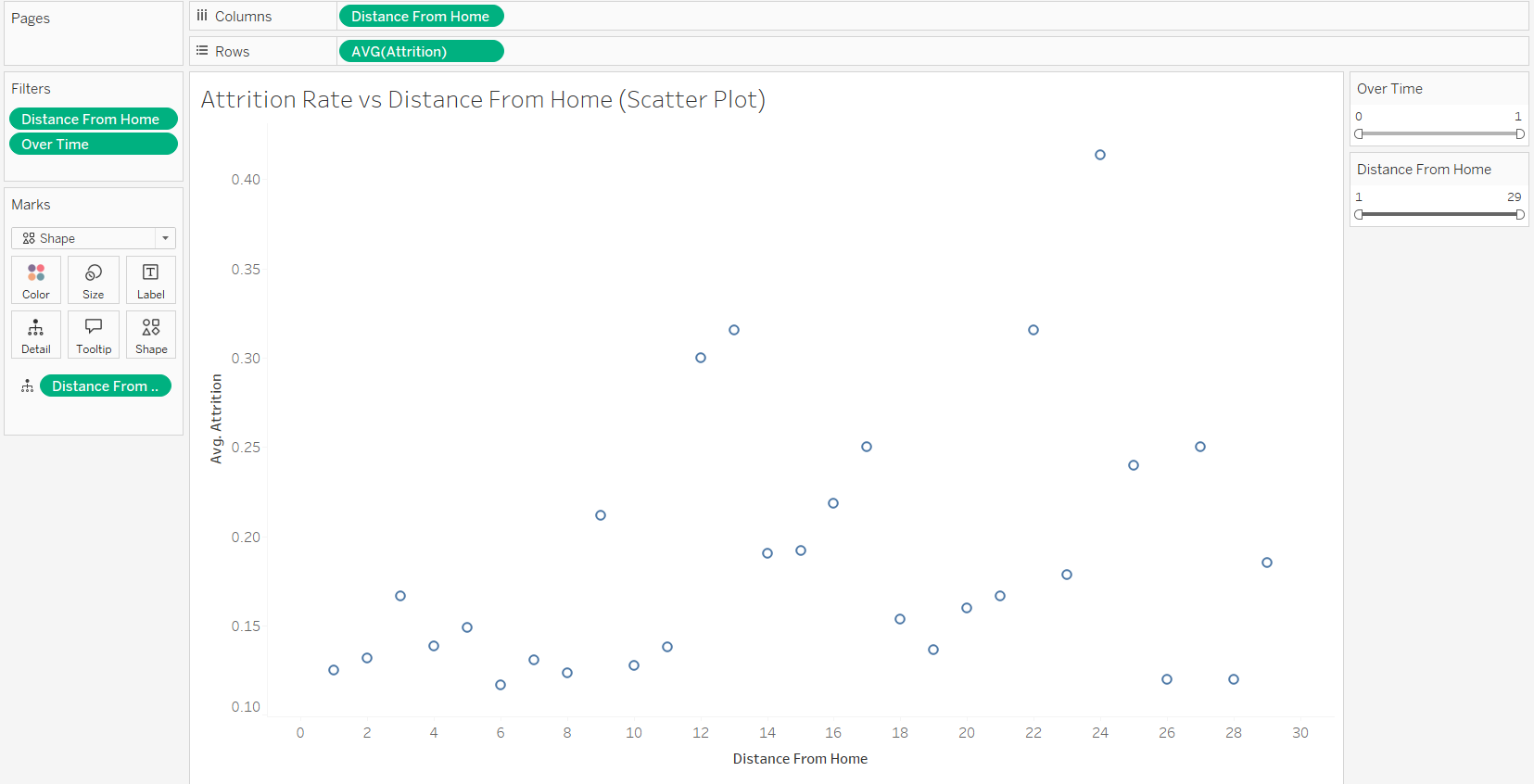
### **What it shows:**

A heatmap where Business Travel categories are on the x-axis and Work-Life Balance ratings are on the y-axis.  
 Color intensity shows the average attrition rate — darker colors represent higher attrition risk.  
 Employees with frequent travel and poor work-life balance show the darkest shades (highest attrition).

### **How it relates:**

This heatmap directly explores how **work pressures like travel demands** interact with **personal satisfaction like work-life balance** to affect attrition rates.  
 It confirms that **heavy travel combined with poor balance significantly increases the chance of leaving**, emphasizing that organizational support for work-life balance is critical to retention.

**Question 11: How does DistanceFromHome correlate with OverTime and Attrition?**



### **Pre-attentive attributes:**

X-Axis = DistanceFromHome

Y-Axis = Avg Attrition

Shape = OverTime (toggle)

Filters: Distance slider, OverTime toggle

### **What it shows:**

A scatterplot where each point represents the average attrition rate at a given distance from home. The shape of the points varies based on OverTime status (Yes/No), helping distinguish patterns of remote workers and overtime employees.

### **How it relates:**

This view highlights whether distance from the workplace and overtime pressures are associated with higher attrition rates. It informs potential remote work or transfer policies for at-risk employees.

**Question 12: How does YearsSinceLastPromotion impact both Attrition risk and PerformanceRating?**



### **Pre-attentive attributes:**

X-Axis = YearsSinceLastPromotion

Y-Axis (left) = Avg Attrition (Line)

Y-Axis (right) = Avg Performance Rating (Bar)

Filter: YearsSinceLastPromotion slider

### **What it shows:**

A dual-axis graph with a line plotting attrition rates and a bar showing average performance ratings over years since last promotion. It allows simultaneous assessment of retention and performance impact.

**How it relates:**

It uncovers if lack of promotion over time is tied to employees leaving or performing differently. This supports HR initiatives targeting timely promotions to boost engagement and minimize turnover.

**Question 13: What’s the distribution of StockOptionLevel, and do richer equity packages boost retention?**



### **Pre-attentive attributes:**

X-Axis = Stock Option Level

Y-Axis = Count of Employees (Bar) and Avg Attrition (Bar)

Filter: StockOptionLevel slider

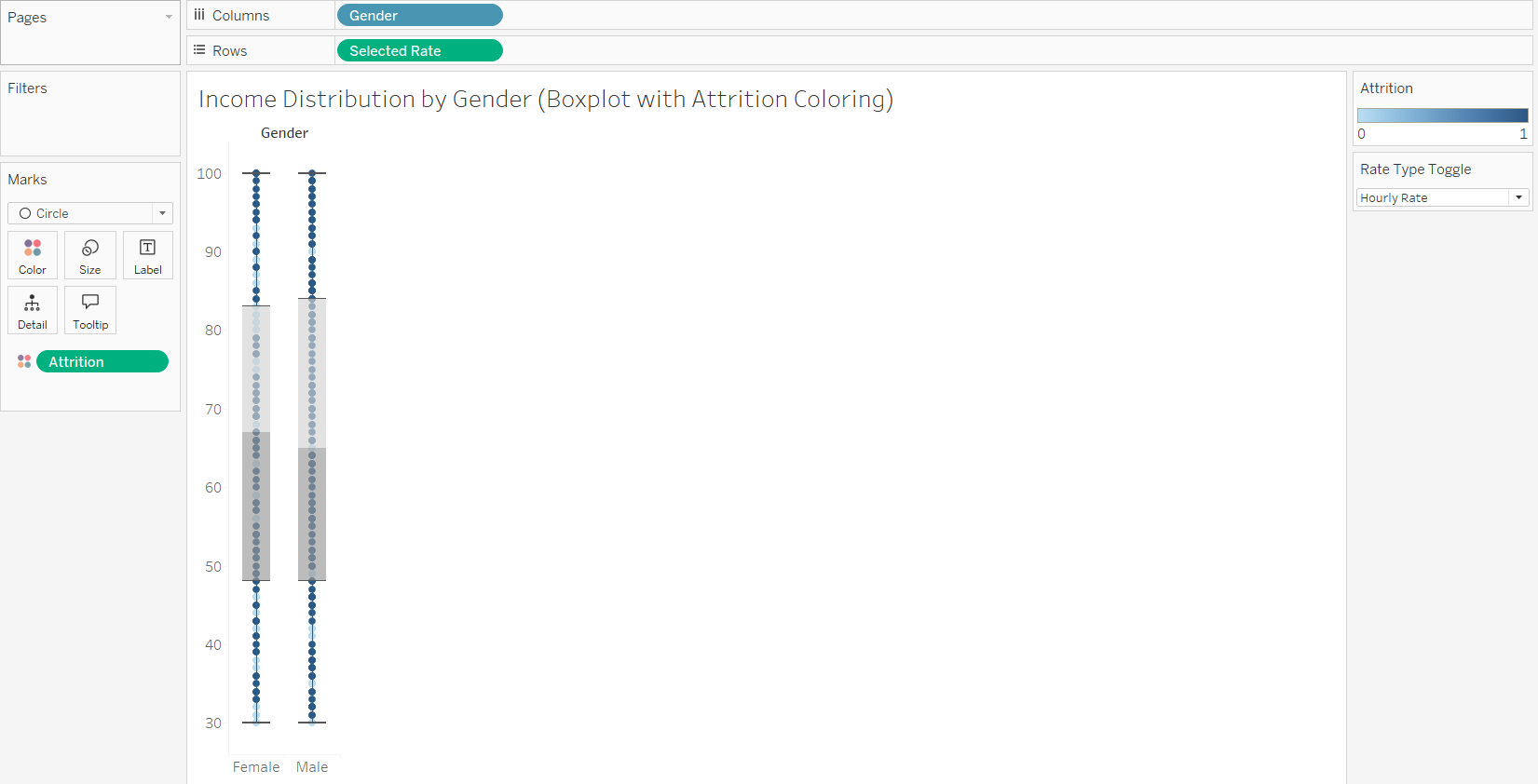
### **What it shows:**

A histogram showing how many employees fall under each StockOptionLevel alongside a bar chart of their respective attrition rates.

**How it relates:**

It explores whether providing better stock options contributes to higher retention, informing equity-based incentive structures.

**Question 14: Are there gender-based differences in pay rates (HourlyRate, DailyRate, MonthlyIncome) that correlate with Attrition?**



### **Pre-attentive attributes:**

X-Axis = Gender

Y-Axis = Selected Pay Rate (toggle between Hourly, Daily, Monthly)

Color = Attrition status

Filters: Gender toggle, Rate Type toggle

### **What it shows:**

A boxplot (or violin plot) showing the distribution of selected pay rates for Male and Female employees, colored by attrition risk.

**How it relates:**

This visualization identifies any potential gender pay gaps and their correlation with turnover risk, assisting diversity and equity policy adjustments.

Dashboard Interactivity

**Question 1: What is our overall attrition rate, and how does it differ by Department and JobRole?**

**Department : Multi-select**

It is loaded from the dataset from Department column itself.

**Job Role: Multi-select**

It is loaded from the dataset from JobRole column itself.

**Question 2: How do EducationFields and Education levels relate to PerformanceRating?**

**Department: Multi-select**

It is loaded from the dataset from Department column itself.

**Question 3: What patterns emerge between PercentSalaryHike and subsequent PerformanceRating?**

**Performance Salary Hike : Slider**

We have salary hike ranging from 11->25, as we took sum of the column, the slider starts with 11 and can go till 25.

**Job Level Updated: MultiSelect**

It is loaded from the dataset from JobLevel Updated column itself.

**Question 4: How do attrition rates vary across AgeGroup and Gender?**

**Controls:**

1. **Gender filter** (multi‐select)  
   * **Use:** Include/exclude male or female cohorts
   * **Connected Plot:** Stacked‐bar chart of Attrition Rate by AgeGroup & Gender
   * **Values:** Loaded from the **Gender** field → {Male, Female}
2. **AgeGroup filter** (multi‐select checklist)  
   * **Use:** Restrict to one or more age buckets
   * **Connected Plot:** Same stacked‐bar chart above
   * **Values:** Loaded from the **AgeGroup** field → {18–25, 26–35, 36–45, 46–55, 55+}

**Question 5: How do the four Satisfaction scores (Environment, Job, Relationship, Work-Life) compare by Department and how do they predict Attrition?**

**Controls:**

1. **Department filter** (dropdown)  
   * **Use:** Slice heatmap to specific functional teams
   * **Connected Plot:** Heatmap of Satisfaction Type × Department (color = avg score, size = attrition)
   * **Values:** Loaded from the **Department** field → e.g. {Sales, R&D, HR, …}
2. **Satisfaction Type filter** (multi‐select)  
   * **Use:** Show/hide one or more of the four survey metrics
   * **Connected Plot:** Same heatmap (facets or pivoted dims)
   * **Values:** Loaded from the **Satisfaction Type** field → {Environment, Job, Relationship, Work–Life}

**Question 6: Which JobLevels and JobRoles show the strongest PerformanceRating and lowest Attrition?**

**Controls:**

1. **Department filter** (dropdown)  
   * **Use:** Limit the bubble‐matrix to a specific department’s roles
   * **Connected Plot:** Bubble‐matrix of JobLevel × JobRole (size = attrition, color = avg performance)
   * **Values:** Loaded from the **Department** field → list of dept names
2. **JobLevel filter** (multi‐select)  
   * **Use:** Focus on one or more seniority tiers
   * **Connected Plot:** Same bubble‐matrix above
   * **Values:** Loaded from the **JobLevel\_Updated** field → {Very Low, Low, Medium, High, Very High}

**Question 7: Is lower JobInvolvement associated with higher Attrition?**

**Controls:**

1. **Department filter** (dropdown)  
   * **Use:** Show attrition‐by‐involvement only for selected teams
   * **Connected Plot:** Bar chart of Attrition Rate by JobInvolvement levels
   * **Values:** Loaded from the **Department** field → list of dept names
2. **JobRole filter** (multi‐select)  
   * **Use:** Drill into specific roles’ involvement‐attrition relationship
   * **Connected Plot:** Same bar chart above
   * **Values:** Loaded from the **JobRole** field → e.g. {Sales Executive, Research Scientist, …}

**Question 8: Is there a relationship between MonthlyIncome (or PercentSalaryHike) and Attrition?**

* **Income Range Slider**
  + **Purpose:** Allows the user to filter employees based on their Monthly Income levels.
  + **Connected Plot:** Scatterplot showing Income, PercentSalaryHike, and Attrition.
  + **Value Range:** $1,009 – $35,722 (based on the MonthlyIncome attribute).
  + **Loaded From:** MonthlyIncome column in the dataset.
* **Percent Salary Hike Slider**
  + **Purpose:** Enables filtering employees based on their Percent Salary Hike.
  + **Connected Plot:** Scatterplot showing Income, PercentSalaryHike, and Attrition.
  + **Value Range:** 11% – 40% (based on the PercentSalaryHike attribute).
  + **Loaded From:** PercentSalaryHike column in the dataset.

**Question 9: Does TrainingTimesLastYear drive higher PerformanceRating?**

* **Training Times Last Year Slider**
  + **Purpose**: Allows the user to filter based on the number of training sessions an employee attended in the last year.
  + **Connected Plot**: Scatterplot showing TrainingTimesLastYear versus PerformanceRating, separated by Department with Polynomial Trend Lines.
  + **Value Range:** 0 – 6 (based on the TrainingTimesLastYear attribute).
  + **Loaded From**: TrainingTimesLastYear column in the dataset.

**Question 10: How do BusinessTravel categories (Rarely, Frequently, None) affect WorkLifeBalance and Attrition?**

* **Business Travel Dropdown**
  + **Purpose:** Allows the user to select and filter by type of Business Travel (e.g., Non-Travel, Travel\_Frequently, Travel\_Rarely).
  + **Connected Plot:** Heatmap showing the relationship between Business Travel, Work-Life Balance, and Attrition Rate.
  + **Value Range:** Non-Travel, Travel\_Frequently, Travel\_Rarely (categorical options from BusinessTravel).
  + **Loaded From:** BusinessTravel column in the dataset.

**Question 11: How does DistanceFromHome correlate with OverTime and Attrition?**

**Controls:**

* DistanceFromHome Range Slider
* OverTime Toggle (Yes/No)

**Purpose:**

* Allows the user to filter employees based on their commuting distance and overtime status.

**Connected Plot:**

* Scatterplot showing DistanceFromHome versus Average Attrition Rate, separated by OverTime status.

**Value Range:**

* DistanceFromHome: 1 – 29 miles
* OverTime: Yes or No (binary)

**Loaded From:**

* DistanceFromHome and OverTime columns in the dataset.

**Question 12: How does YearsSinceLastPromotion impact both Attrition risk and PerformanceRating?**

**Controls:**

* YearsSinceLastPromotion Range Slider

**Purpose:**

* Allows the user to filter employees based on the number of years since their last promotion.

**Connected Plot:**

* Dual-axis chart showing Attrition Rate (line) and Average Performance Rating (bar) against Years Since Last Promotion.

**Value Range:**

* 0 – 15 years (based on YearsSinceLastPromotion attribute).

**Loaded From:**

* YearsSinceLastPromotion column in the dataset.

**Question 13: What’s the distribution of StockOptionLevel, and do richer equity packages boost retention?**

**Controls:**

* StockOptionLevel Dropdown or Slider

**Purpose:**

* Allows the user to select and filter employees based on their Stock Option Level.

**Connected Plot:**

* Histogram showing number of employees by Stock Option Level and a bar chart showing Attrition Rate.

**Value Range:**

* 0, 1, 2, 3 (discrete levels).

**Loaded From:**

* StockOptionLevel column in the dataset.

### **Question 14: Are there gender-based differences in pay rates (HourlyRate, DailyRate, MonthlyIncome) that correlate with Attrition?**

**Controls:**

* Gender Toggle (Male/Female)
* Pay Rate Type Dropdown (HourlyRate, DailyRate, MonthlyIncome)

**Purpose:**

* Allows the user to filter employees by Gender and choose among different Pay Rate types.

**Connected Plot:**

* Boxplot showing distribution of selected Pay Rate by Gender, colored by Attrition status.

**Value Range:**

* Gender: Male, Female
* Pay Rate Types: HourlyRate, DailyRate, MonthlyIncome

**Loaded From:**

* Gender, HourlyRate, DailyRate, and MonthlyIncome columns in the dataset.

References

Mural Board Link: <https://app.mural.co/t/iftdv533projectgroup159816/m/iftdv533projectgroup159816/1744753409465/d76d1cd50c961e0798b26fc4e201db71a6c6fd82>