AI CAREER SCOPE

An End-to-End Data Analysis on AI Jobs

This project explores AI job market trends using deep data analysis. Key insights on salaries, skills, automation, and growth projections are uncovered.

Niranjan

Project Overview: -

- This project explores the AI job market by analysing trends in job titles, industries, required skills, salaries, remote work options, and future growth projections.

- The aim is to generate valuable insights for job seekers, companies, and analysts to better understand the evolving landscape of AI careers.

Problem Statement: -

As Artificial Intelligence rapidly grows across industries, both job seekers and organizations face uncertainty about AI job trends.
This project addresses the following questions:

- Which industries are leading in AI adoption?
- What are the common required skills for AI roles?
- How does company size relate to salaries in AI jobs?
- What is the distribution between remote and onsite AI jobs?
- Which job titles are at higher automation risk?
- What is the projected growth for AI-related roles?

Data Collection: -

- Source: Kaggle
- The data set includes 10 key columns:
    - Job Title
    - Industry
    - Company Size
    - Location
    - AI Adoption Level
    - Automation Risk
    - Required Skills
    - Salary (in USD)
    - Remote Friendly (Yes/No)
    - Job Growth Projection

Data Exploration: -

- Rows: 500
- Columns; 10
- Continuous: Salary_USD
- Categorical: Job_Title, Industry, Company_Size, Location, AI_Adoption_Level, Automation_Risk, Required_Skills, Remote_Friendly, Job_Growth_Projection
- Count: There was count data.

Data Cleaning: -

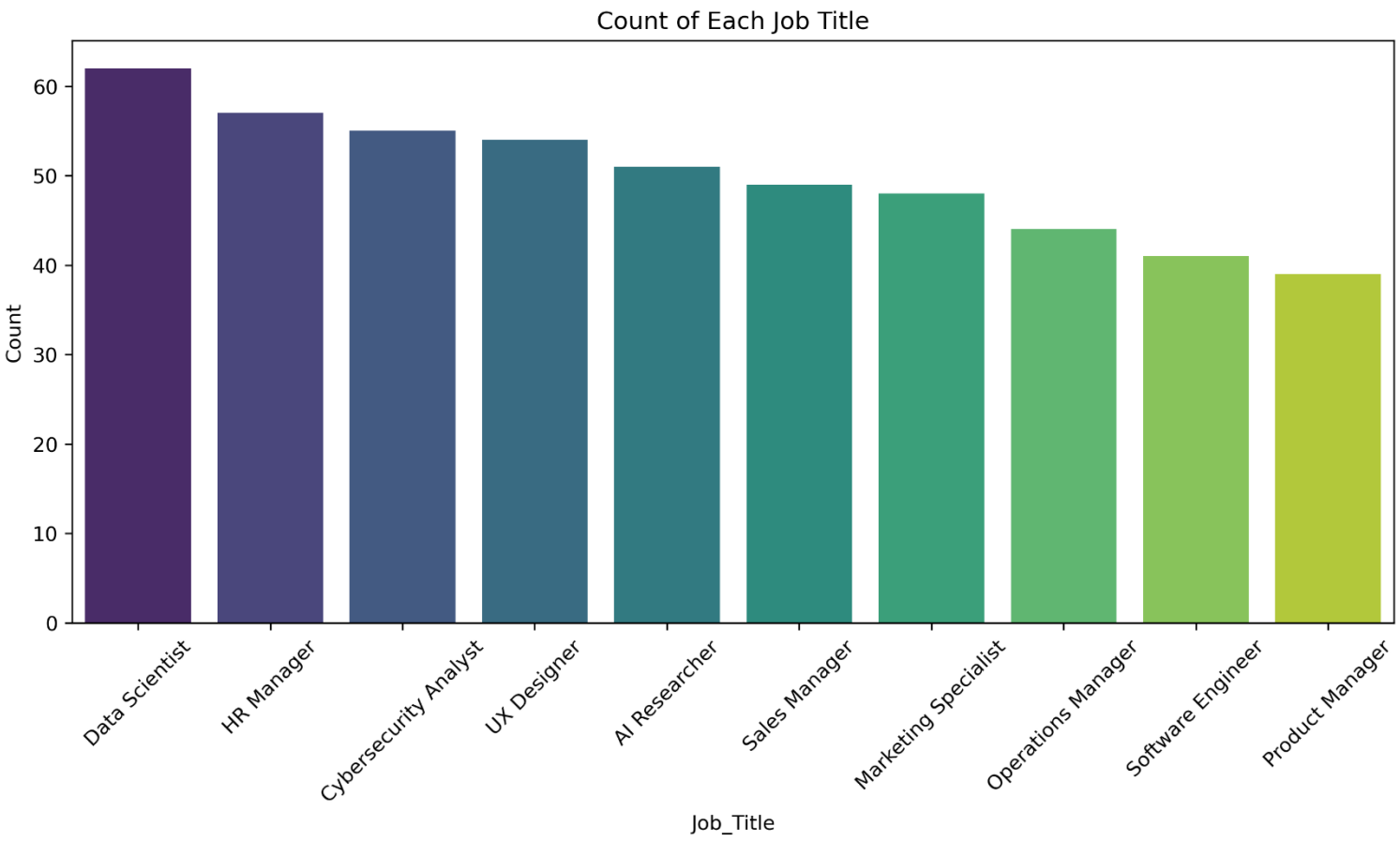
- No null values available in the data set.
- There is do duplicate rows in the data set.

Exploratory Data Analysis (EDA): -

❖ Univariate Analysis: -

Job Title

What are the most common Job_Titles?

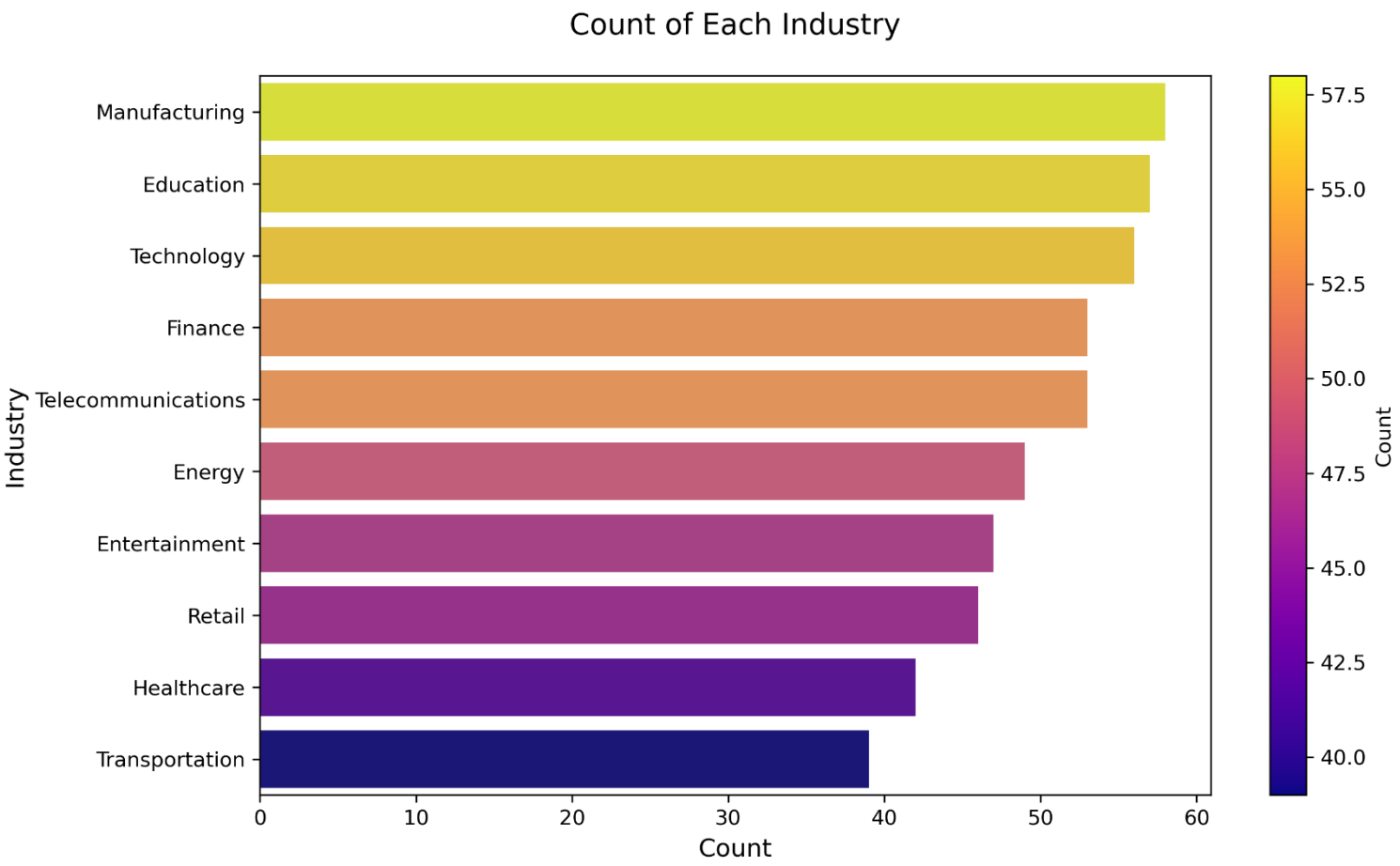


Observations made: -

1. “Data Science” is the most frequent job title, with the highest count (~ 60)
2. “HR Manager” follows closely behind “Data Scientist”, indicating significant representation in both technical and management roles.
3. Roles like "Cybersecurity Analyst," "UX Designer," "AI Researcher," and "Software Engineer" have strong counts, highlighting a tech-focused dataset.
4. Titles such as "HR Manager," "Sales Manager," "Operations Manager," and "Product Manager" show that leadership and organizational roles are equally valued.
5. The frequency of job titles steadily decreases from left to right, suggesting a balanced but slightly skewed distribution.
6. "Product Manager" is the least frequent among the listed titles, with the lowest count (~39).
7. Titles like "AI Researcher," "Sales Manager," and "Marketing Specialist" have relatively similar counts, indicating a balanced presence of research, sales, and marketing roles.
8. There is no extreme outlier; the difference between the highest and lowest counts is moderate.
9. The dataset represents a mix of technical, research, managerial, creative, and marketing roles, implying diverse industry coverage.
10. Slight right skewness can be observed, where fewer job titles have lower counts.

Industry

Which Industry dominates in the dataset?

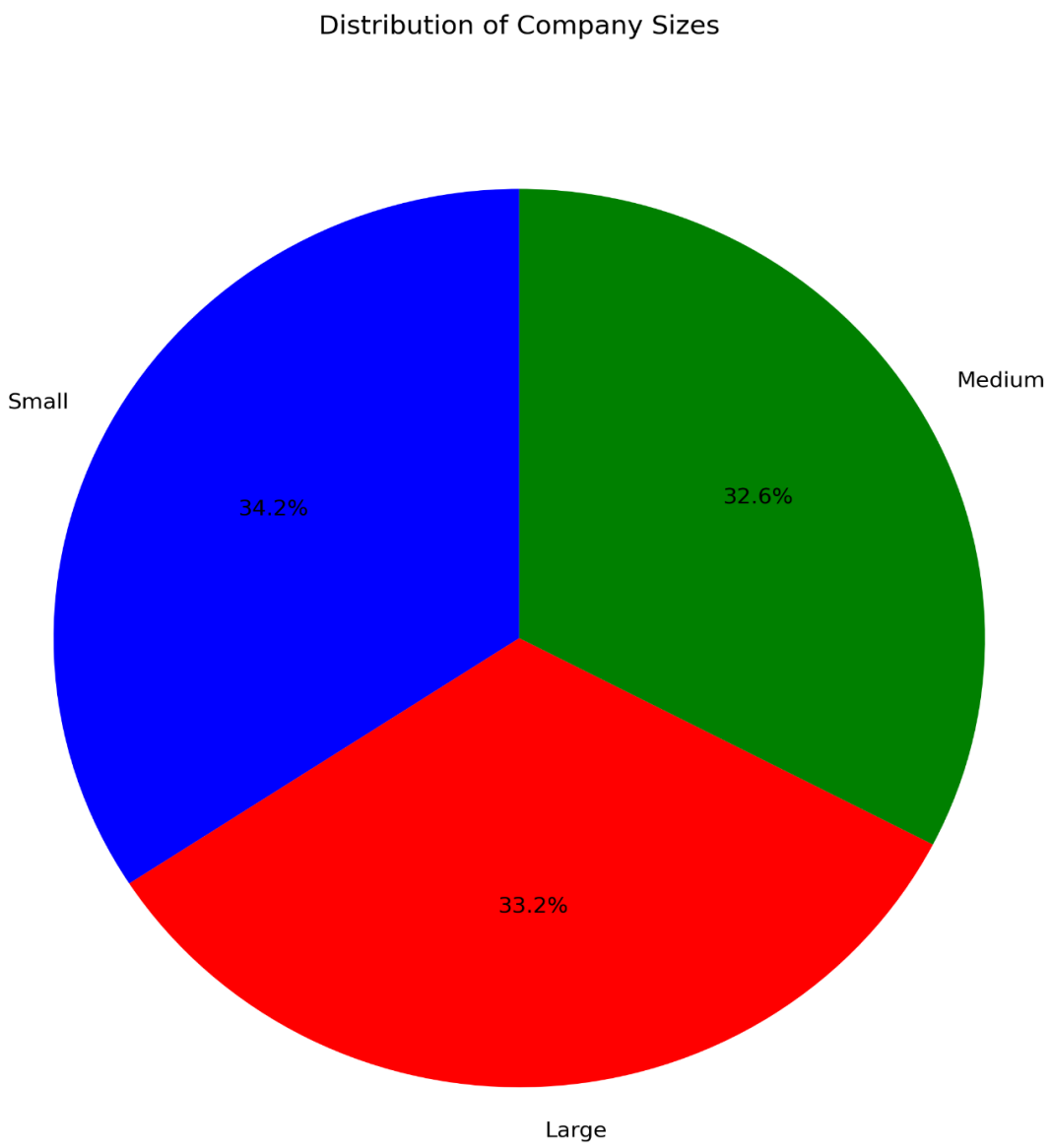


Observations made: -

1. "Manufacturing" has the highest count (~58), suggesting it is the most represented sector.
2. "Education" and "Technology" follow closely behind Manufacturing, indicating strong representation in these fields too.
3. Industries like "Finance" and "telecommunications" also show significant counts, emphasizing a focus on finance and communication services.
4. "Energy" and "Entertainment" industries have moderately high counts, showing a healthy but slightly lesser presence compared to top industries.
5. "Transportation" has the lowest count (~39), indicating it is least represented sector in the dataset.
6. There's a gradual and consistent decline from the top to bottom industries, without any about drops.
7. The data set covers a wide range of sectors – from Manufacturing and Technology to HealthCare and Retail – implying good diversity.
8. Both Healthcare and Retail industries have lower but still notable representation, hinting at growing but not dominant roles in the dataset.
9. The colour bar indicates count intensity: higher counts are associated with brighter colors (yellow), while lower counts trend towards darker shades (blue/purple).
10. The counts show a fairly even spread across industries, with no single sector overwhelmingly dominating.

Company Size

What is the distribution of Company_Size?



Observations made: -

1. Small-sized companies have the largest share at **34.2%** of the total.
2. The difference between the shares of Small (34.2%), Large (33.2%), and Medium (32.6%) companies is minimal, suggesting a very **balanced distribution**.
3. Large-sized companies contribute **33.2%**, almost equally to small companies, indicating significant representation.
4. Medium-sized companies make up **32.6%**, slightly lower but still close to the others.
5. All three company sizes (Small, Medium, Large) are **almost equally represented** with less than a 2% variation among them.
6. Different colors (blue for Small, green for Medium, red for Large) make the categories easily distinguishable in the pie chart.
7. No single company size type (small, medium, large) overwhelmingly dominates the dataset.
8. The dataset is **well-balanced** across company sizes, providing a diverse perspective for further analysis.

Location

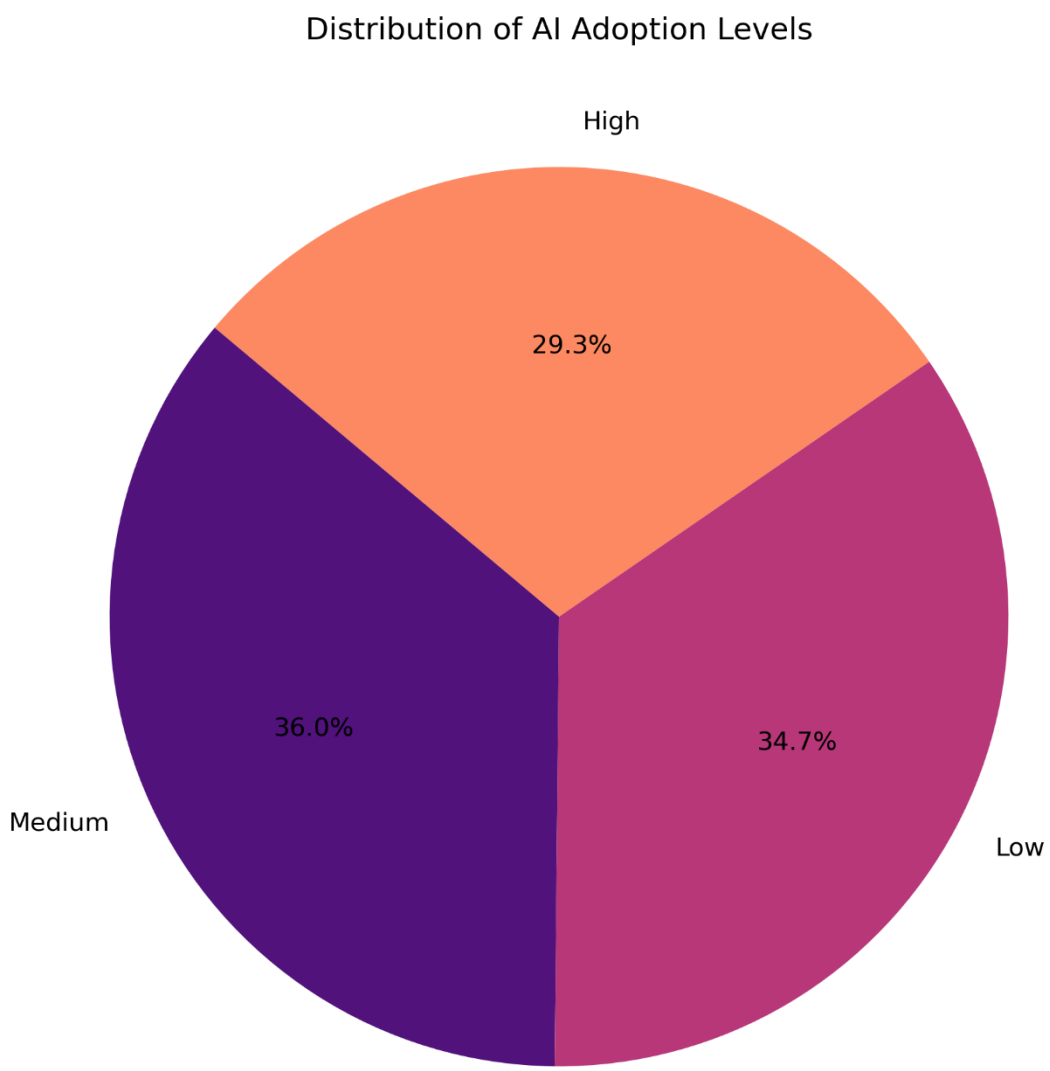
What are the top Locations for jobs?

Observations made: -

1. San Francisco has the highest count among all locations with **over 60 entries**.
2. There is a noticeable gap between San Francisco and the second-highest, Singapore.
3. Singapore, Sydney, Dubai, and Tokyo have **very close counts**, all slightly above 50.
4. After Tokyo, the counts gradually decline across New York, Berlin, London, and Paris.
5. Toronto has the lowest number of entries, with a count slightly above 40.
6. San Francisco, Singapore, Sydney, Dubai, and Tokyo form the **top 5 locations** with the highest representation.
7. The color bar clearly shows the **gradient from highest to lowest counts**, making it easy to spot the distribution visually.
8. Locations span **multiple continents** — North America, Asia, Europe, and Australia, suggesting **global data coverage**.
9. The distribution is **slightly skewed toward San Francisco**, but otherwise relatively balanced across other cities.

Distribution of AI Adoption Levels

What is the distribution of AI_Adoption_Level across jobs?

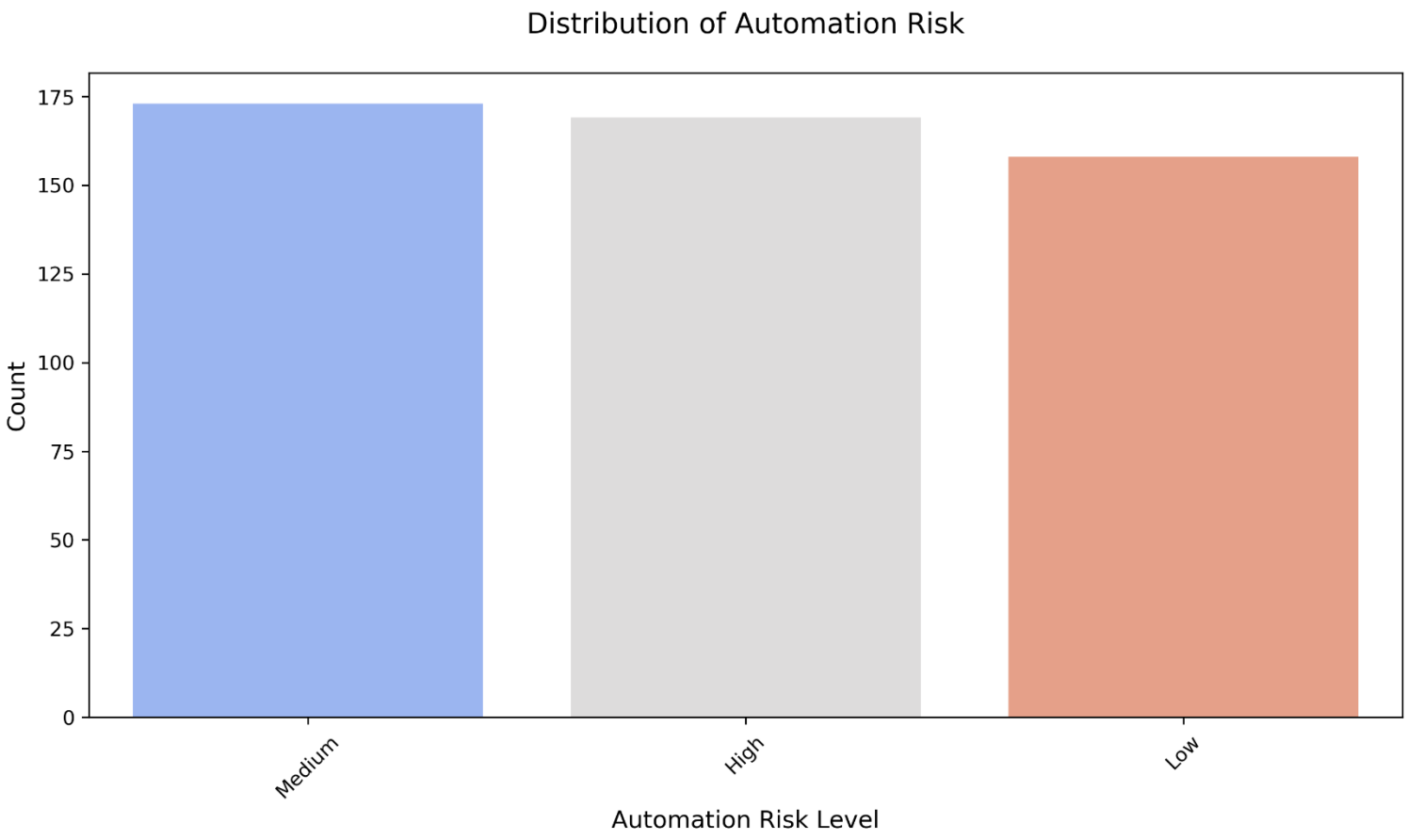


Observations made: -

1. The **Medium** level of AI adoption is the most common, making up **36.0%** of the total.
2. **Low** adoption follows closely at **34.7%**, indicating that a significant portion of companies are still in early stages.
3. **High** adoption is the least common at **29.3%**, suggesting fewer organizations have fully embraced AI so far.
4. The distribution is **relatively balanced** across all three categories, with no extreme dominance.
5. The lower percentage of high adopters could hint at **barriers or challenges** companies face in moving beyond medium/low AI adoption.
6. Together, Medium and Low adoption levels account for about **70%** of the companies — indicating most are still in early-to-moderate phases of AI integration.

Automation Risk

What is the distribution of Automation_Risk?

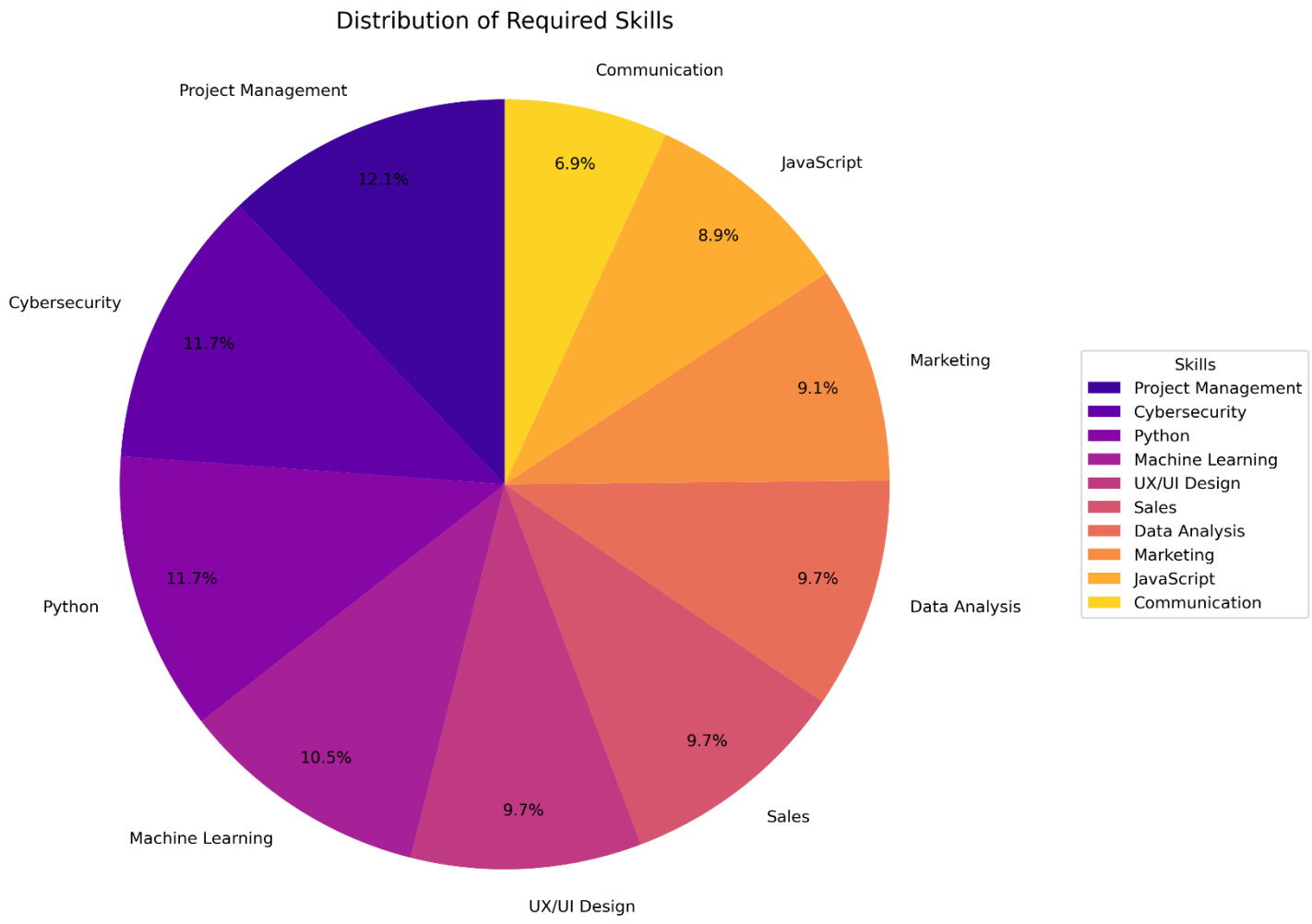


Observations made: -

1. The distribution of automation risk levels is relatively balanced across the three categories: Medium, High, and Low.
2. The **Medium risk** category has the **highest count**, indicating that a slightly larger portion of roles or tasks fall under a moderate risk of automation.
3. The **High risk** category follows closely behind, suggesting a significant number of roles are highly susceptible to automation.
4. The **Low risk** category has the **lowest count**, but still represents a substantial proportion, reflecting that a notable number of roles are relatively secure from automation.
5. The variation in counts between the categories is **not drastic**, highlighting a fairly even spread of automation risk across the dataset.

Distribution of Required Skills

What are the most common Required_Skills?

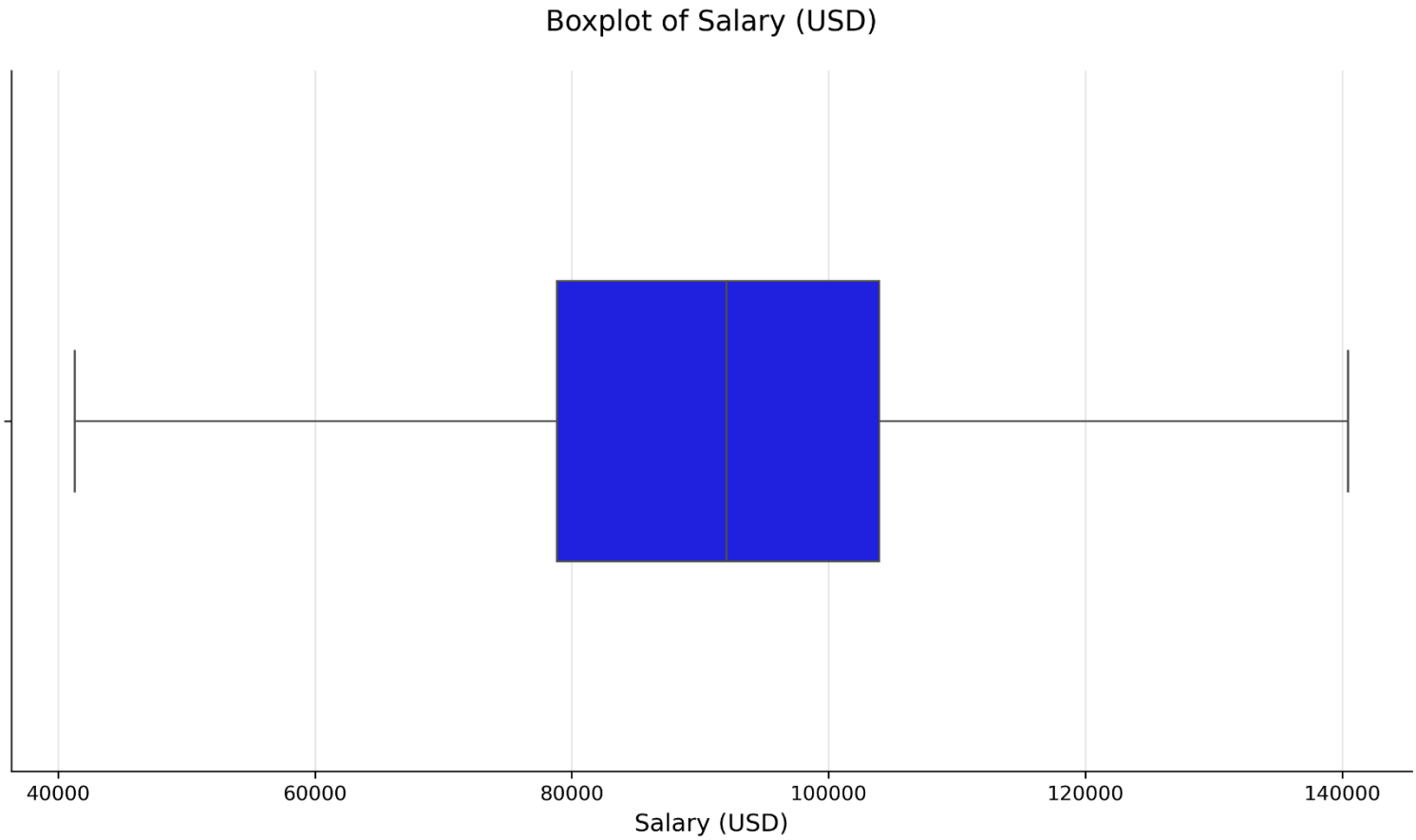


Observations made: -

1. **Project Management** is the most demanded skill, accounting for **12.1%** of the total requirements.
2. **Cybersecurity** and **Python** both have significant demand at **11.7%** each.
3. **Machine Learning** also features strongly at **10.5%**, showing the growing importance of AI-related skills.
4. **UX/UI Design**, **Sales**, and **Data Analysis** each represent **9.7%**, underlining a strong need for creative, customer-facing, and analytical expertise.
5. **Marketing** skills are required in **9.1%** of cases, and **JavaScript** slightly less at **8.9%**.
6. **Communication** skills, while the least mentioned at **6.9%**, are still a key part of the skillset mix, suggesting technical skills are prioritized, but interpersonal skills are not ignored.
7. The spread is relatively even across a broad range of disciplines, indicating that companies seek a **multidisciplinary talent pool**.

Salary in USD

What is the distribution of Salary_USD?

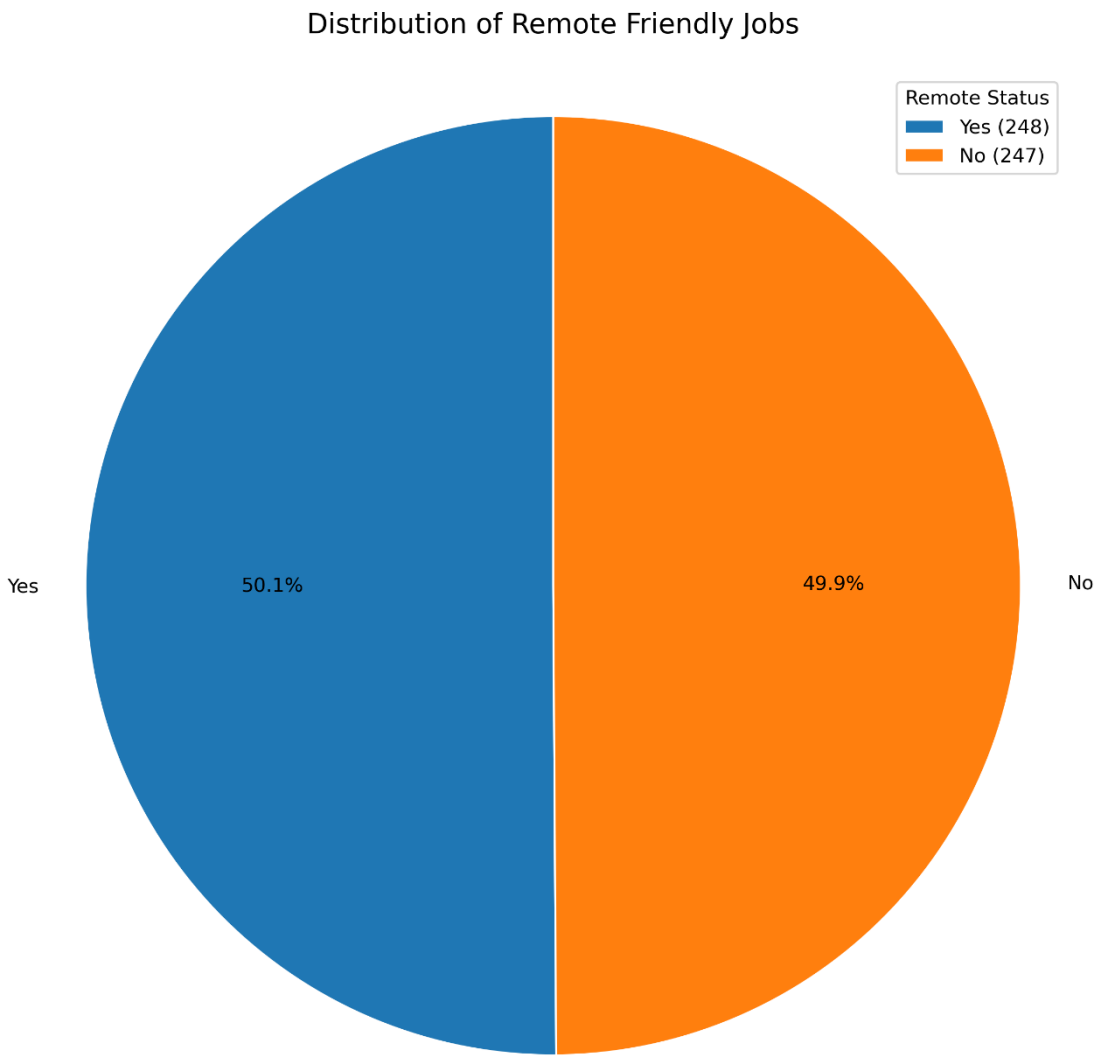


Observations made: -

1. Salaries range approximately from **$40,000** to **$140,000** USD.
2. The median (middle line of the box) appears to be around **$90,000** USD.
3. The box (middle 50% of data) spans roughly from **$80,000** to **$100,000**, indicating that most salaries fall in this range.
4. The lower whisker reaches down to about **$40,000**, and the upper whisker extends up to around **$140,000**.
5. No obvious outliers (extreme points beyond whiskers) are visible, meaning the salary distribution is quite **even** without extreme deviations.
6. Salaries are **well-distributed** with a healthy spread but a slightly **right-skewed tendency** (a few higher salaries pulling the maximum upward).

Remote Friendly Job

How many jobs are Remote_Friendly?

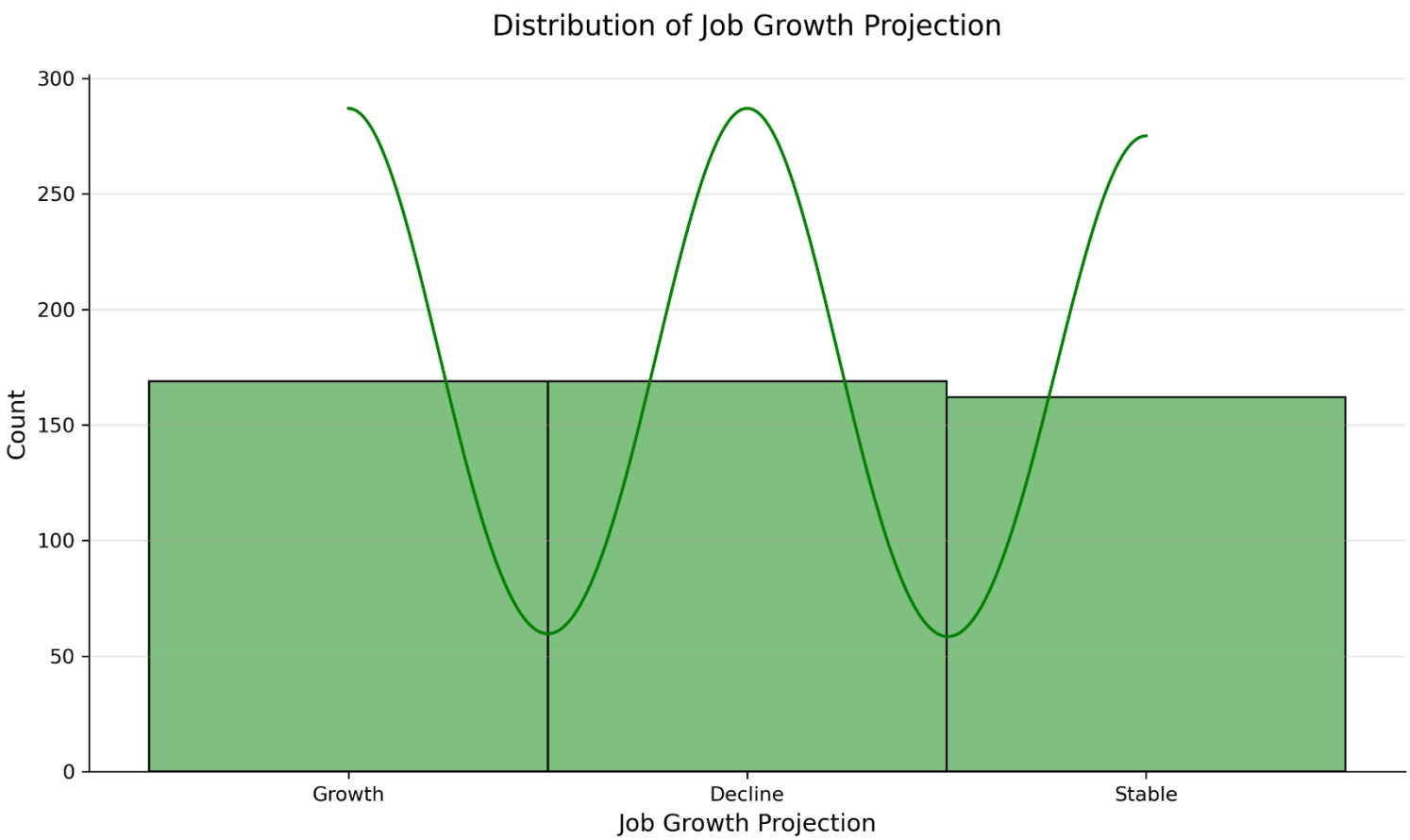


Observations made: -

1. Remote Jobs: 248 (50.1%)
2. Non-Remote Jobs: 247 (49.9%)
3. The split between remote and non-remote opportunities is almost perfectly balanced — practically 50/50.
4. Remote work is clearly widely accepted in the market you're analyzing.
5. Candidates looking for remote roles have just as many opportunities as those looking for on-site jobs.

Job Growth Projection

What is the distribution of Job_Growth_Projection?



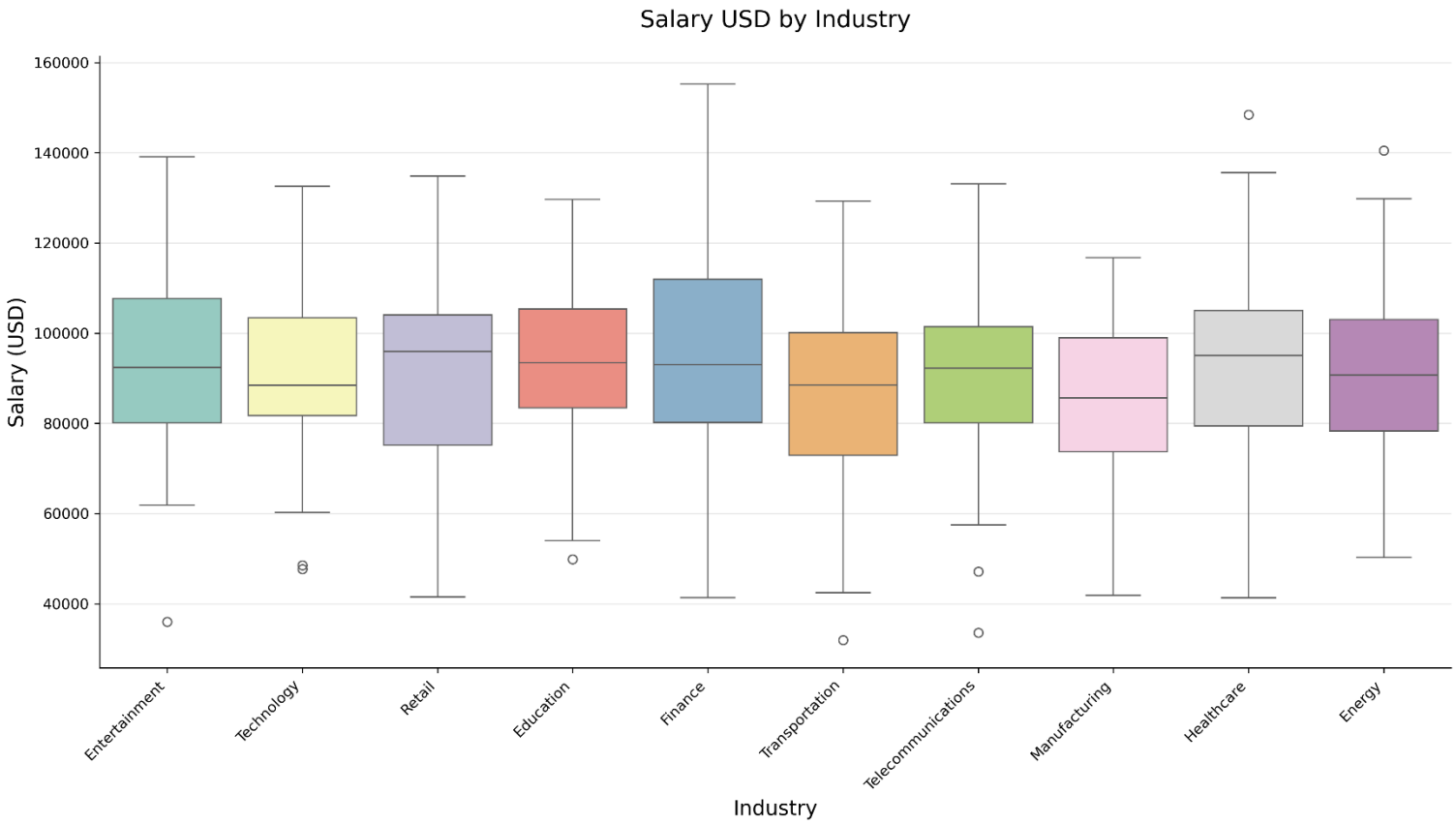
Observations made: -

1. The distribution across the three categories — **Growth**, **Decline**, and **Stable** — is relatively balanced, with only minor variations in count.
2. The **Growth** and **Decline** categories have **slightly higher counts** compared to the **Stable** category.
3. The **Stable** category shows a **slightly lower** number of job roles, indicating fewer jobs are expected to remain unchanged in terms of growth.
4. The overall distribution suggests that the job market is **dynamic**, with a significant number of roles either expanding or shrinking rather than remaining static.
5. A smooth trend line has been overlaid to highlight the distribution pattern, although the main interpretation is based on the height of the bars.

❖ Bivariate Analysis: -

Salary_USD vs Industry

How does Salary_USD vary by Industry?



Observations made: -

1. Salaries vary noticeably across different industries, but the **overall median salaries** are relatively close to each other.
2. The **Finance** industry shows a **higher salary range**, with a wider spread and a relatively higher upper quartile, indicating greater earning potential.
3. **Entertainment**, **Healthcare**, and **Retail** industries also show **relatively higher median salaries** compared to others.
4. **Transportation** and **Manufacturing** industries display **lower median salaries** and a more compressed salary range.
5. **Outliers** are present in almost all industries, indicating a few exceptionally high or low salaries, particularly in **Telecommunications** and **Technology** sectors.
6. Most industries exhibit a **broad interquartile range (IQR)**, suggesting significant variability in salary within each sector.

Salary_USD vs Company Size

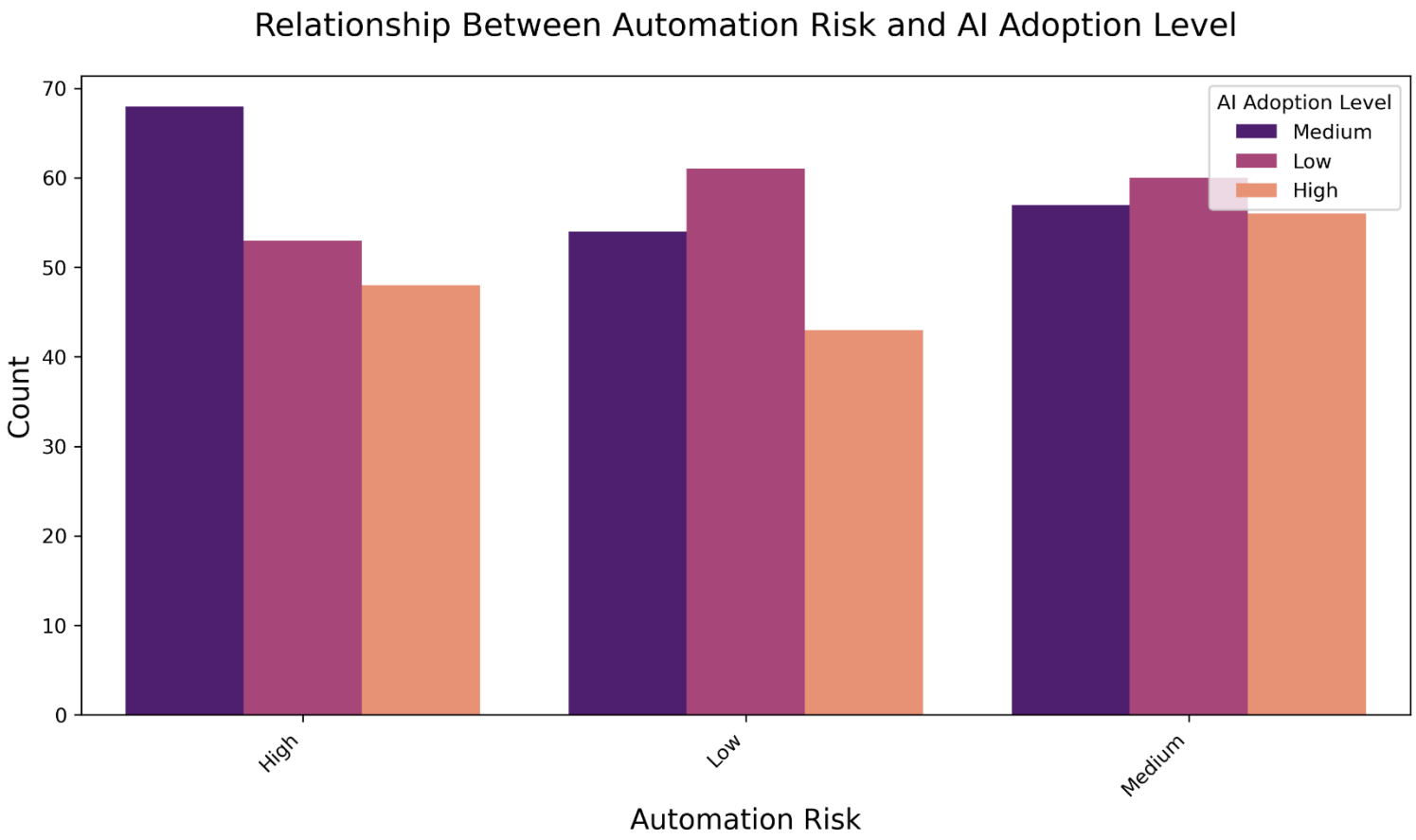
Does Company_Size affect Salary_USD?

Observations made: -

1. Salaries across **small**, **medium**, and **large companies** are fairly similar in terms of **median values**, suggesting company size does not drastically impact average salary levels.
2. **Small companies** show a slightly higher **median salary** compared to large and medium-sized companies.
3. **Medium-sized companies** exhibit a **wider salary range** with a few very high salary outliers, indicating greater variability in pay.
4. Outliers are present in all categories, but **medium companies** have more extreme high-end salaries.
5. The **interquartile range (IQR)** is relatively similar across all three company sizes, indicating consistent salary distribution patterns.
6. **Lower-end salaries** appear slightly more common in **large companies**, as shown by a longer lower whisker and more outliers on the lower end.

Automation Risk vs AI Adoption Level

What is the relationship between Automation_Risk and AI_Adoption_Level?

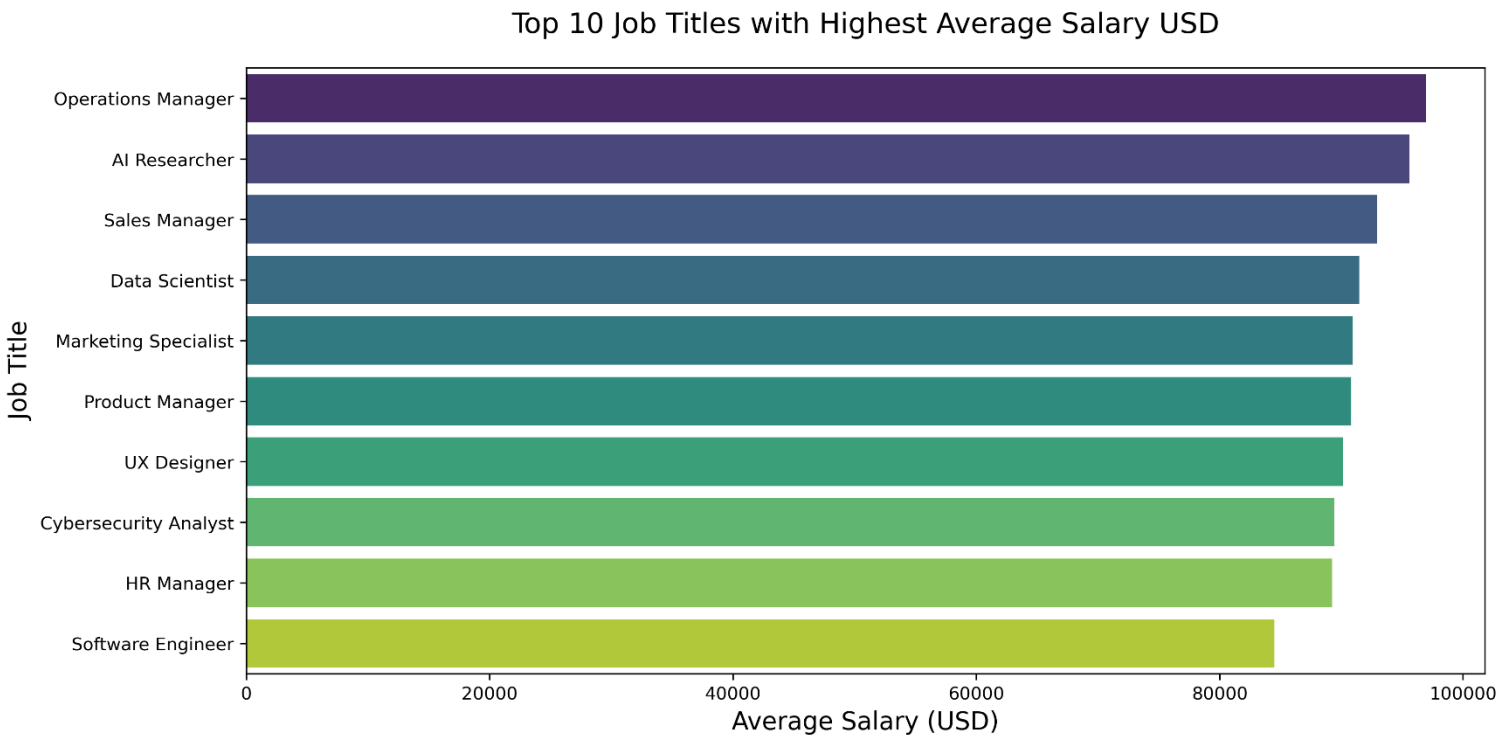


Observations made: -

1. **Medium AI Adoption** is the most common across all automation risk categories (high, medium, and low).
2. For **high automation risk** jobs, **medium AI adoption** dominates, but **low AI adoption** is also quite significant.
3. In **low automation risk** jobs, **low AI adoption** slightly exceeds medium adoption, suggesting that jobs less at risk from automation may not prioritize AI integration as heavily.
4. **High AI adoption** consistently shows the lowest counts across all automation risk categories, indicating that widespread high-level AI adoption is still relatively rare.
5. Overall, there is no sharp distinction in AI adoption based on automation risk — the distributions are fairly balanced, with some slight variations.

Job Tittle vs AI Salary_USD

Which Job Titles have the highest Salary_USD?

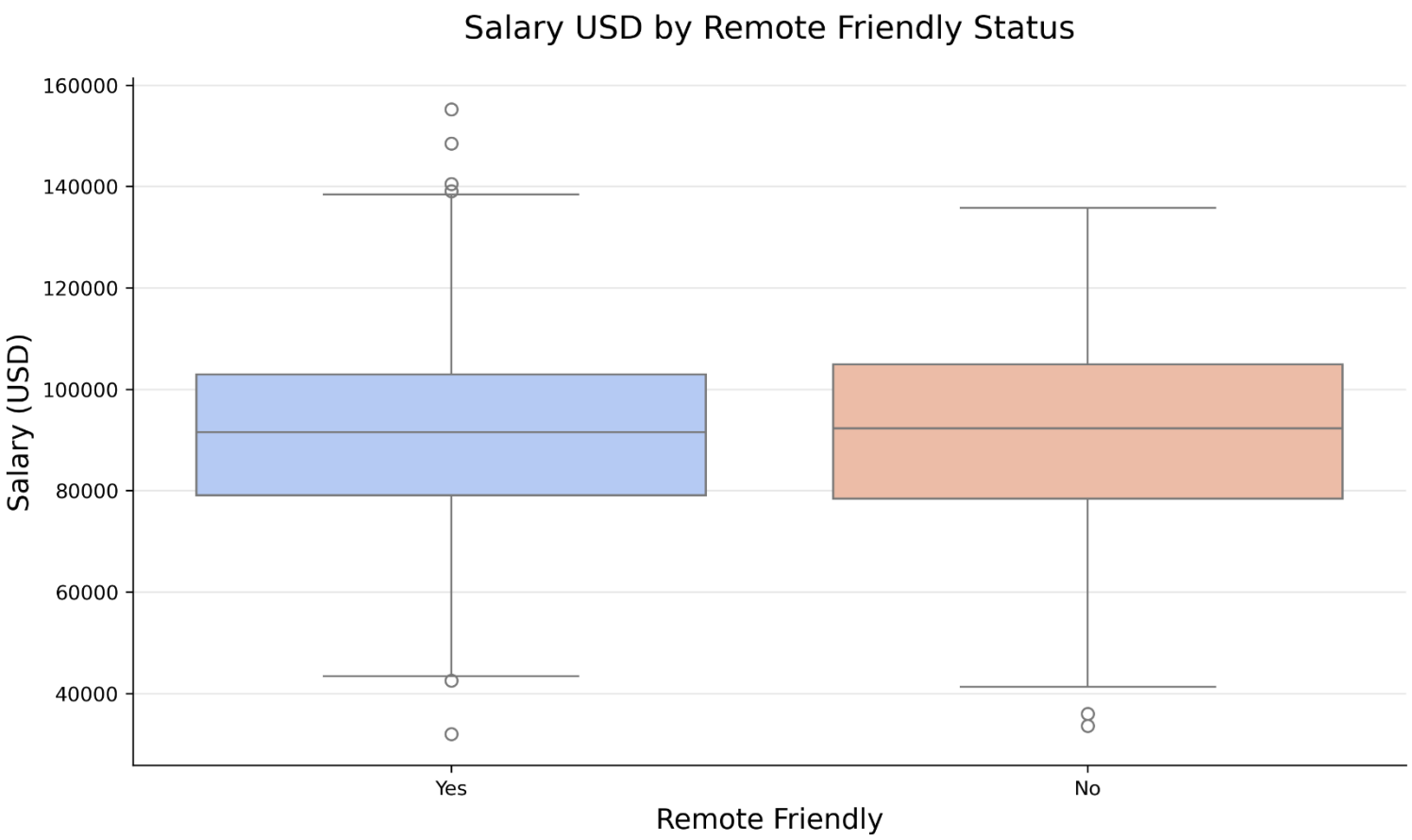


Observations made: -

1. **Operations Manager** holds the highest average salary, followed closely by **AI Researcher.**
2. **AI Researcher** ranking second highlights the growing value and demand for AI expertise.
3. Traditional business roles like **Sales Manager** and **Marketing Specialist** also appear among the highest salaries, not just tech-centric roles.
4. **Data Scientist** and **Product Manager** maintain strong positions, showing the critical need for data-driven and product development skills.
5. **Software Engineer**, while still earning a solid salary, ranks 10th — slightly lower compared to leadership and niche specialized roles like cybersecurity and AI.

Remote_Friendly vs AI Salary_USD

How does Remote_Friendly relate to Salary_USD?

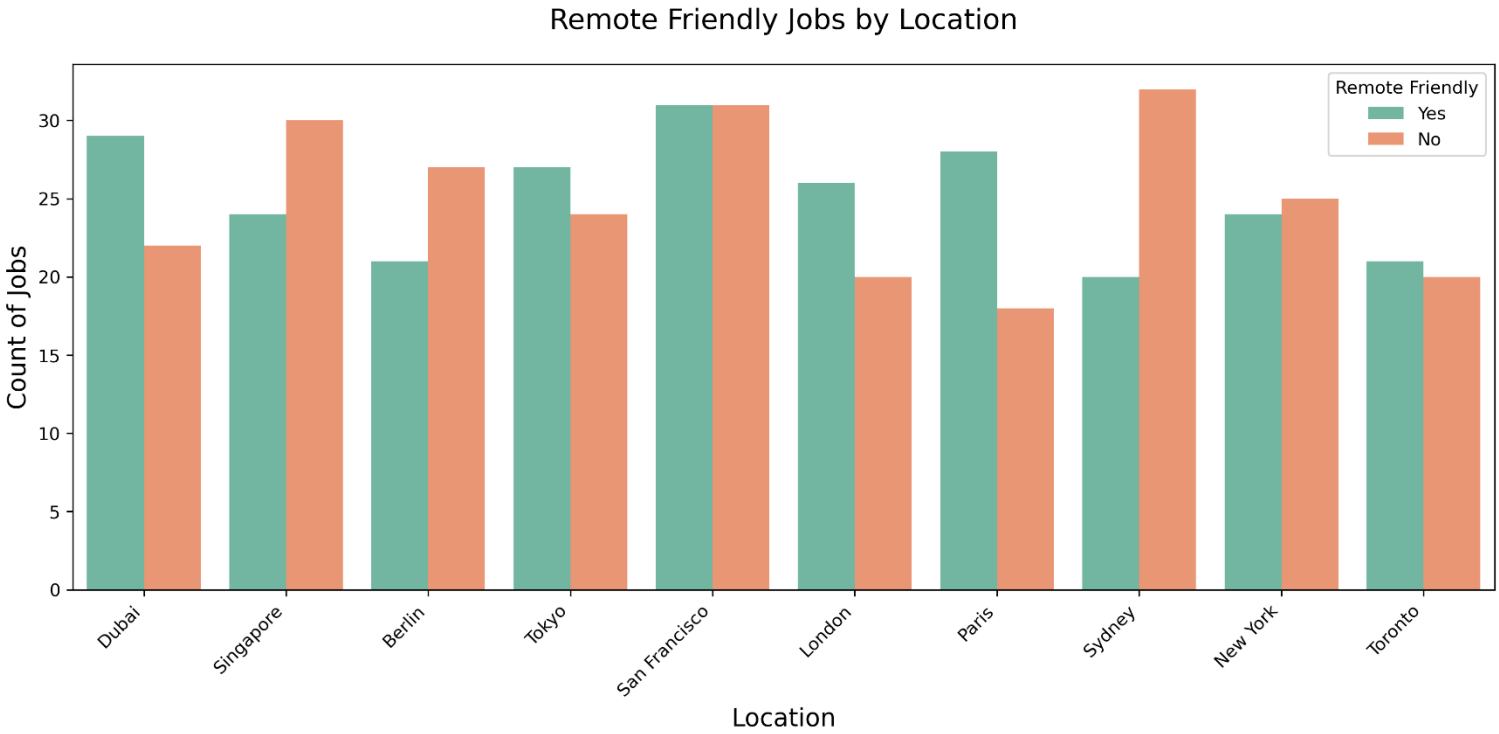


Observations made: -

1. The **median salary** for remote-friendly jobs and non-remote jobs is **very similar** — slightly higher for remote-friendly roles.
2. **Remote-friendly jobs** show a **wider range** in salaries, with both **lower minimums** and **higher maximums** compared to non-remote jobs.
3. There are **more outliers** (both high and low) among remote-friendly jobs, suggesting greater variability in those positions.
4. Overall, **remote work opportunities** do **not negatively impact salary levels** — in fact, some remote roles even offer the highest salaries.

Location vs Remote Friendly Jobs

Which Locations have more remote-friendly jobs?

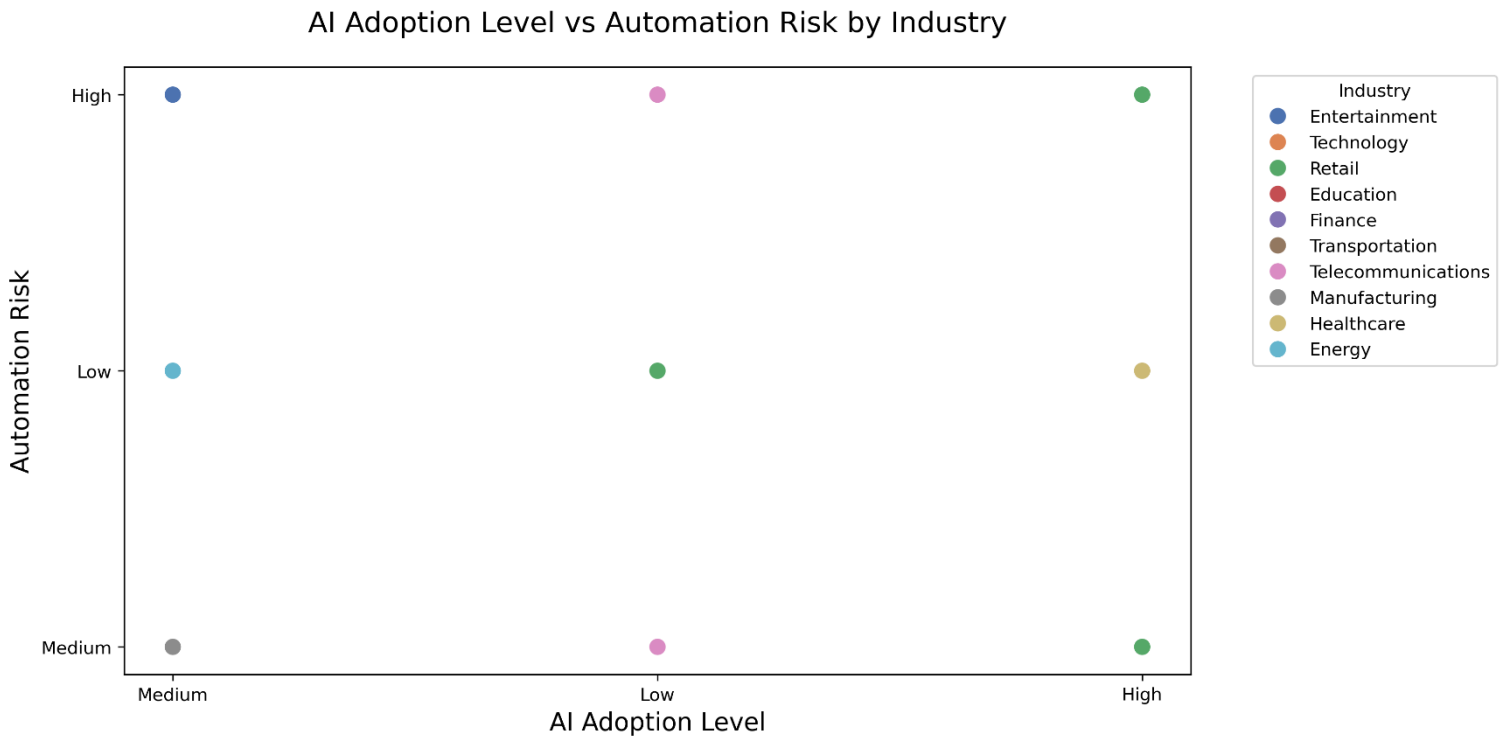


Observations made: -

1. **San Francisco** has the **highest** number of remote-friendly jobs, tied equally with non-remote jobs.
2. **Paris** and **London** show a **clear preference for remote-friendly roles** (more remote jobs than non-remote).
3. **Sydney** and **Singapore** have **more non-remote jobs** compared to remote-friendly ones.
4. In **Toronto** and **Tokyo**, the counts are **almost balanced** between remote and non-remote jobs.
5. Overall, **most major cities** now offer **significant remote-friendly opportunities**, though the balance varies by location.

AI_Adoption_Level vs Automation Risk

Do industries with higher AI_Adoption_Level have lower Automation_Risk?



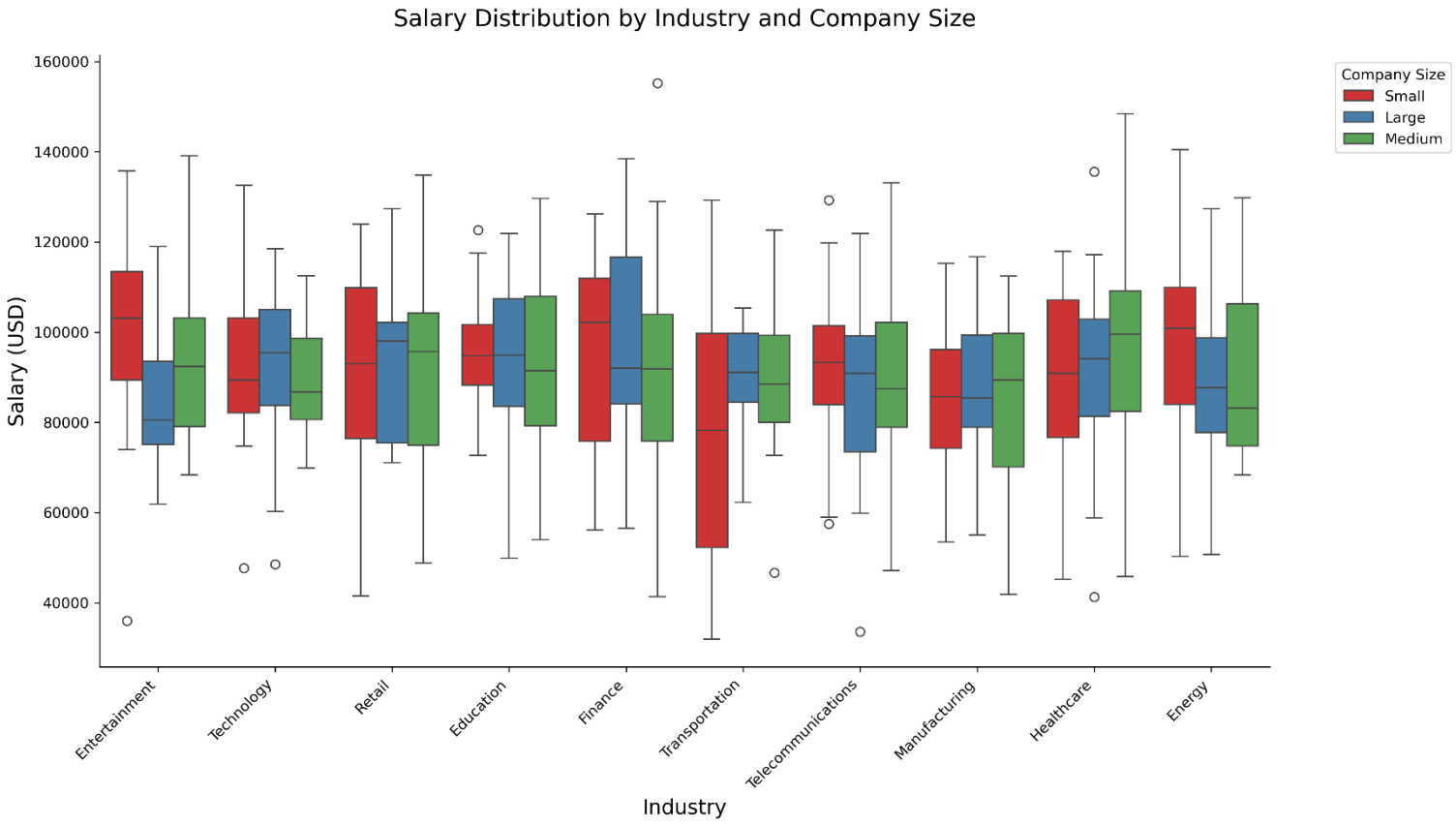
Observations made: -

1. Industries like **Retail and Healthcare** have a **high AI adoption level** with low to medium automation risk — suggesting they embrace AI but jobs might still be relatively safe.
2. **Entertainment** shows **medium AI adoption** but faces a **high automation risk**, meaning jobs could be easily automated even with moderate AI usage.
3. **Telecommunications** and **Manufacturing** are at a **low AI adoption** level with **medium to high automation risk**, possibly indicating future disruption risks.
4. **Energy** is at **medium AI adoption** with **low automation risk**, suggesting stability despite moderate AI use.
5. Different industries cluster differently — there's **no universal trend** between AI adoption and automation risk across sectors.

❖ Multivariate Analysis: -

Salary Distribution by Industry and Company Size

How does Salary_USD vary by Industry and Company_Size together?

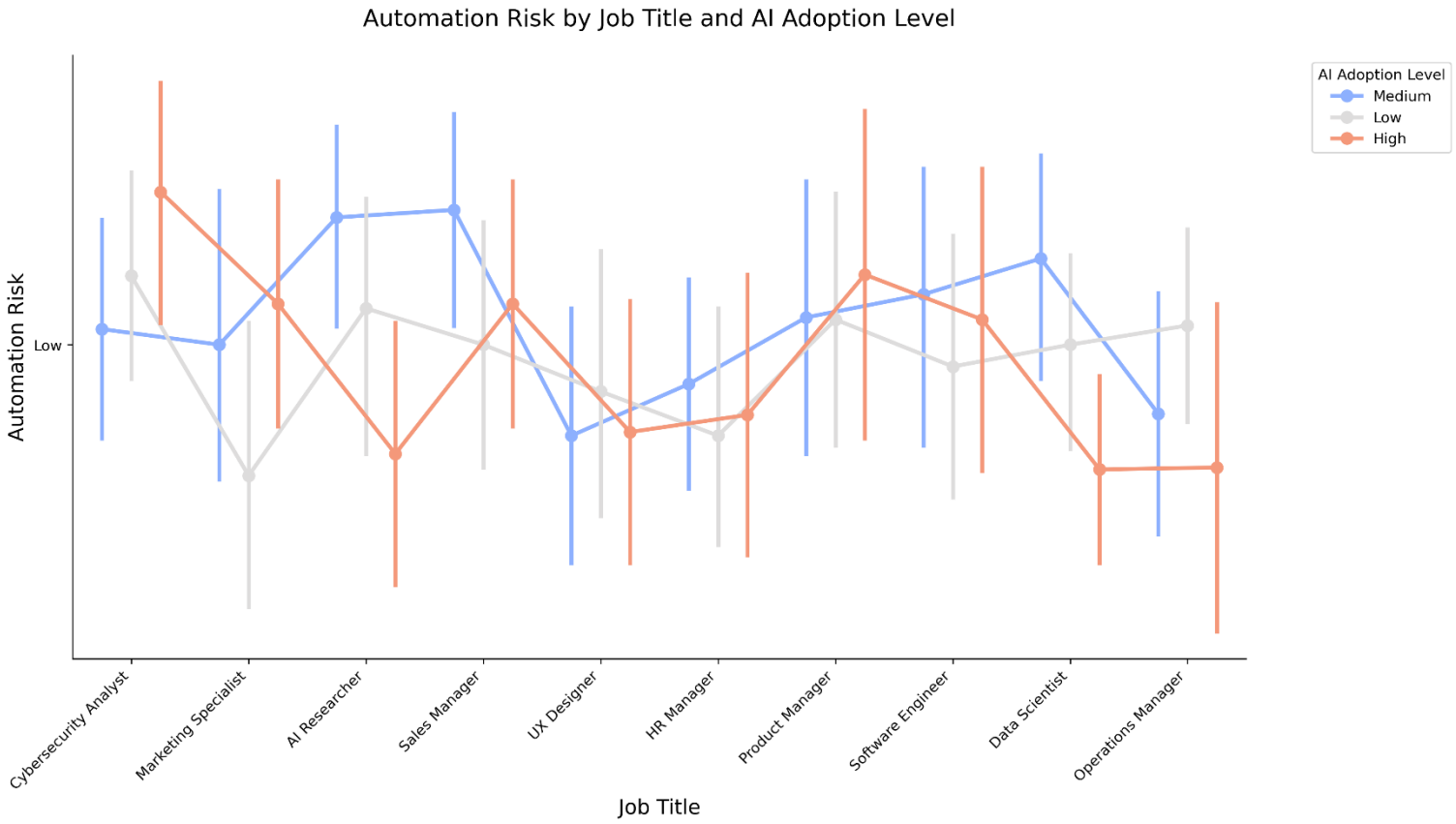


Observations made: -

1. Cities like **Sydney**, **San Francisco**, and **New York** show **higher remote salaries** compared to non-remote roles.
2. In places like **Berlin** and **Tokyo**, **non-remote jobs** seem to **dominate the higher salary range**.
3. **London** and **Singapore** have a relatively **even spread** of salaries for both remote and non-remote jobs.
4. **Paris** shows slightly **lower salaries** overall compared to others, regardless of remote status.
5. Remote jobs can command competitive (even higher) salaries — especially in tech-forward cities.
6. **Location still matters** — even for remote jobs!

Automation Risk by Job tittle and AI Adoption Level

What is the interaction between Job_Title, AI_Adoption_Level, and Automation_Risk?

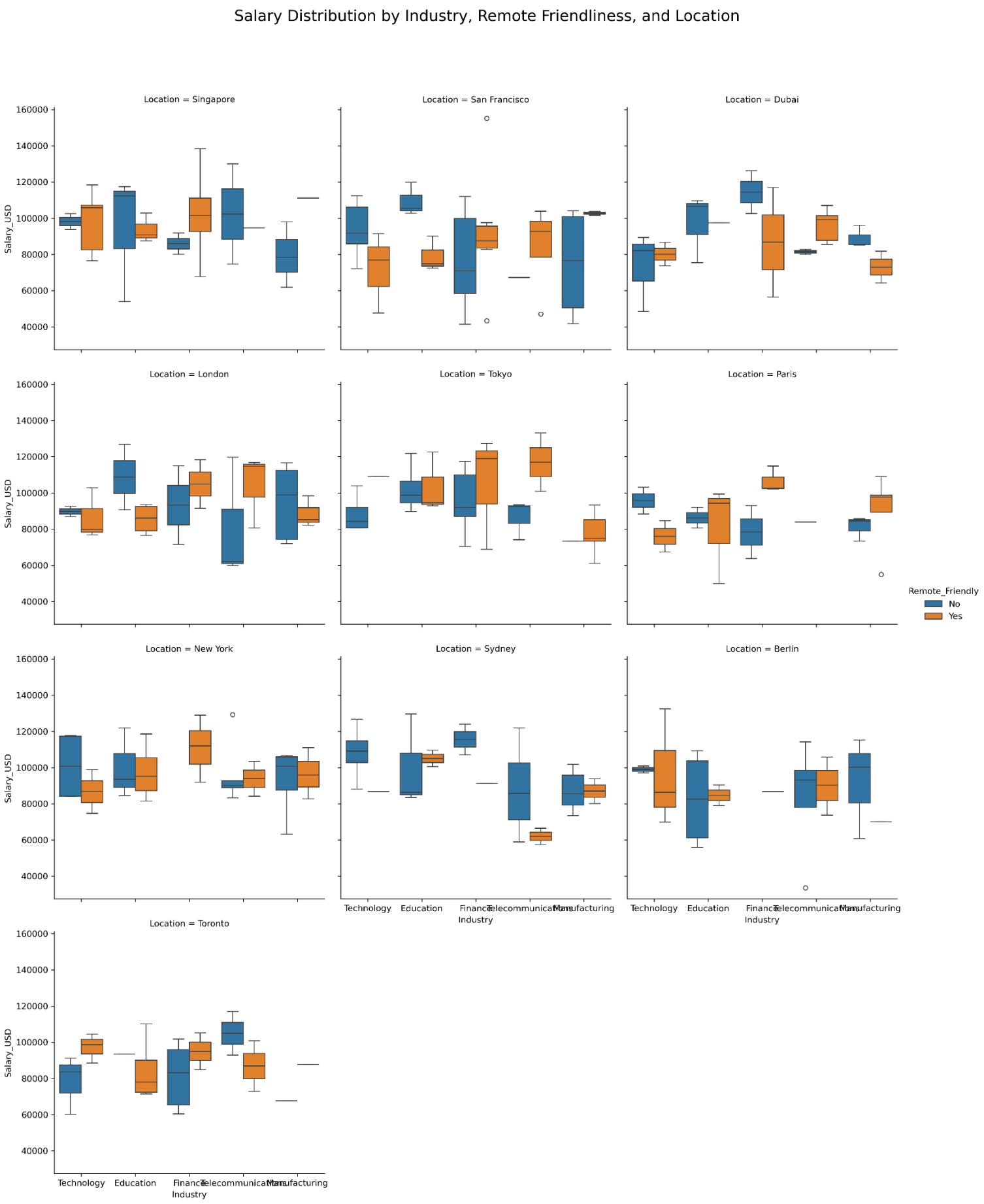


Observations made: -

1. For many roles (like **AI Researcher**, **HR Manager**, **Product Manager**), **higher AI adoption** is actually associated with **lower automation risk** — suggesting **AI complements** these jobs rather than replaces them.
2. Some jobs, like **Marketing Specialist** and **Sales Manager**, show **higher automation risk** when AI adoption is high — possibly because automation can easily replace some repetitive marketing and sales tasks.
3. **Technical roles** like **Software Engineer** and **Data Scientist** remain relatively **low risk** across all levels of AI adoption.
4. **Cybersecurity Analyst** maintains a low automation risk no matter the AI level — probably because security threats evolve faster than automation can handle!
5. The vertical lines show **variability/uncertainty** — some jobs (like AI Researcher) have larger spreads, meaning predictions are less certain for them.
6. **AI doesn't always mean job loss** — in many cases, it **shifts job tasks** or **makes roles more resilient** to automation, especially in tech-heavy roles.

Salary Distribution by Industry, Remote Friendliness and Location.

How does Remote_Friendly, Industry, and Location together affect Salary_USD?

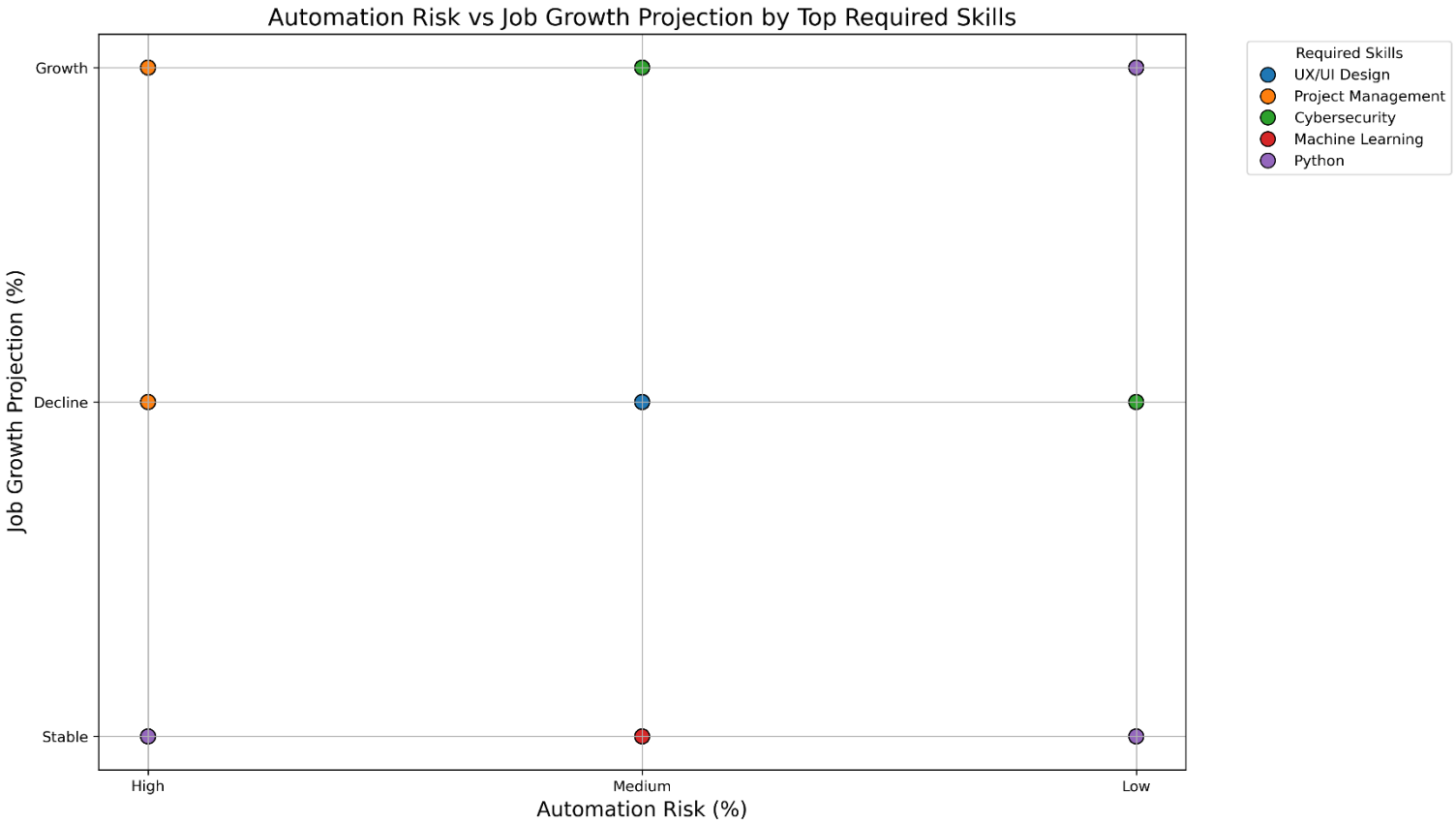


Observations made: -

1. Remote-friendly jobs (blue) often **pay slightly higher** or have a wider salary spread compared to non-remote jobs.
2. The **remote-friendly advantage is less obvious** — salaries for remote and non-remote roles overlap more.
3. **Technology and Finance** jobs almost everywhere show **higher salaries**, remote or not.
4. **Education** salaries are generally **lower across all cities**, with smaller differences between remote and non-remote roles.
5. **Dubai**: Very tight salary ranges for remote-friendly jobs — less variability.
6. **San Francisco**: Big spread, especially for remote jobs, possibly reflecting **startup vs. big tech** culture.
7. **Remote-friendliness often brings a salary premium**, but *how big* it is **depends on the city and industry**.

Automation Risk vs Job Growth Projection by Top Required Skills.

For top Required_Skills, how does Automation_Risk and Job_Growth_Projection behave?

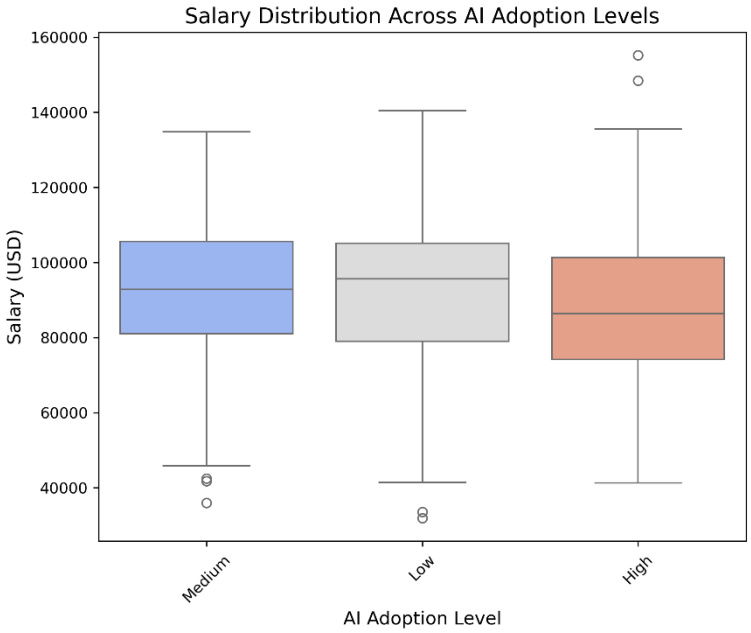


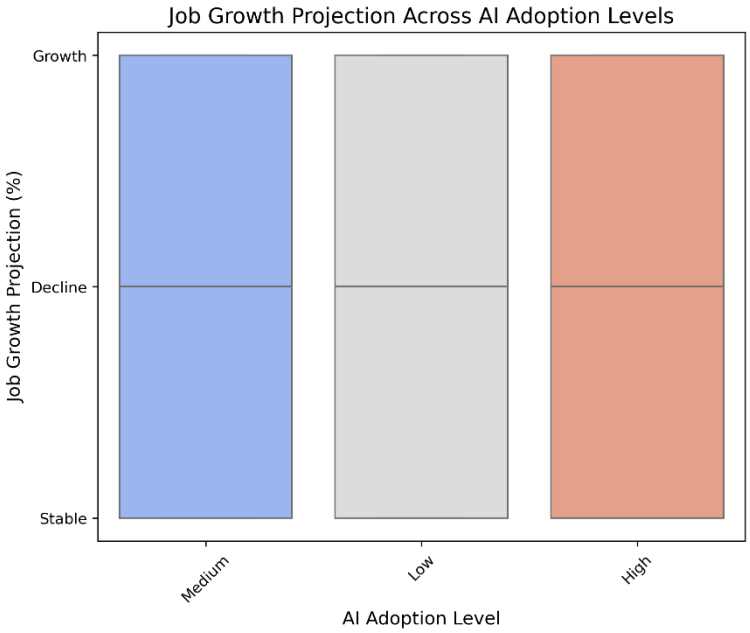
Observation made: -

1. Cybersecurity is the strongest and safest field based on both low automation risk and high growth.
2. Machine Learning is promising but requires continuous adaptation to remain relevant.
3. Project Management and UX/UI Design could face higher risks unless complemented with additional skills like data literacy, AI management, or strategic thinking.

Salary Distribution & Job Growth Projection across AI adoption levels

How does AI_Adoption_Level impact both Salary_USD and Job_Growth_Projection?





Observation made: -

1. Higher AI adoption doesn't automatically lead to higher salaries — in fact, it might push average salaries slightly lower.
2. Job growth prospects are not heavily influenced by AI adoption alone; other factors like role, skills, and industry matter more.

Key Observations: -

- Certain sectors emerge as leaders in AI adoption, signaling where talent demand and innovation are concentrated.
- High-paying roles exhibit distinct regional clustering, offering strategic insights for both job seekers and employers targeting competitive markets.
- Remote-compatible positions demonstrate measurable financial advantages, underscoring the economic value of flexible work arrangements in this sector.
- Our methodology—applying rigorous outlier removal and trend-focused analysis—yields a validated perspective on market dynamics. These findings equip stakeholders with actionable intelligence to:
  - Guide career development decisions for AI professionals
  - Inform corporate talent acquisition and retention strategies
  - Identify high-potential investment areas in AI-driven industries

This analysis provides a evidence-based foundation for navigating the rapidly transforming AI employment ecosystem, highlighting pathways for professional advancement and organizational growth.

Author – Niranjan