<u>AI Career Scope</u>

Author: Niranjan

Date: 1 April 2025

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.

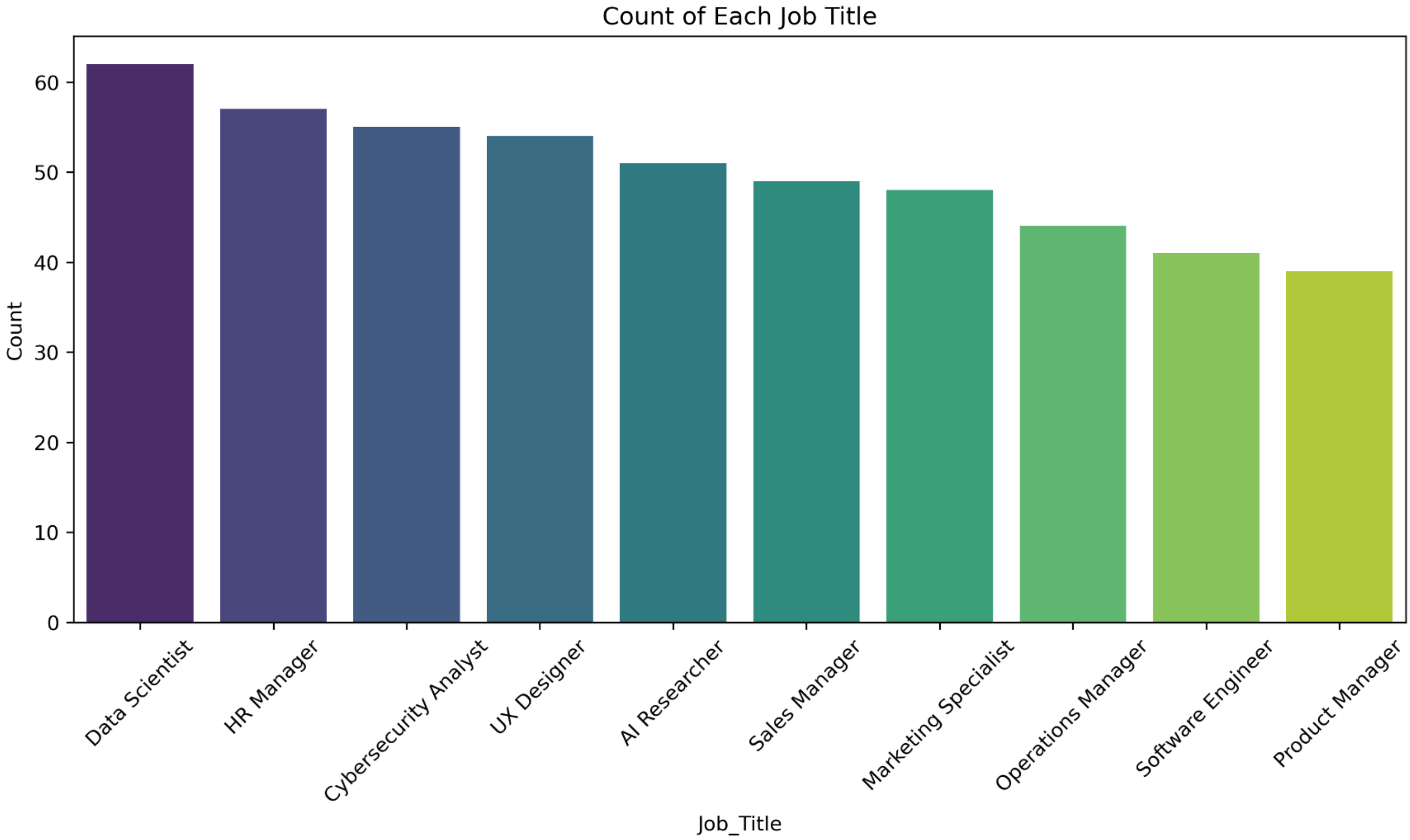
Data Cleaning: -

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

Exploratory Data Analysis (EDA): -

❖ Univariate Analysis: -

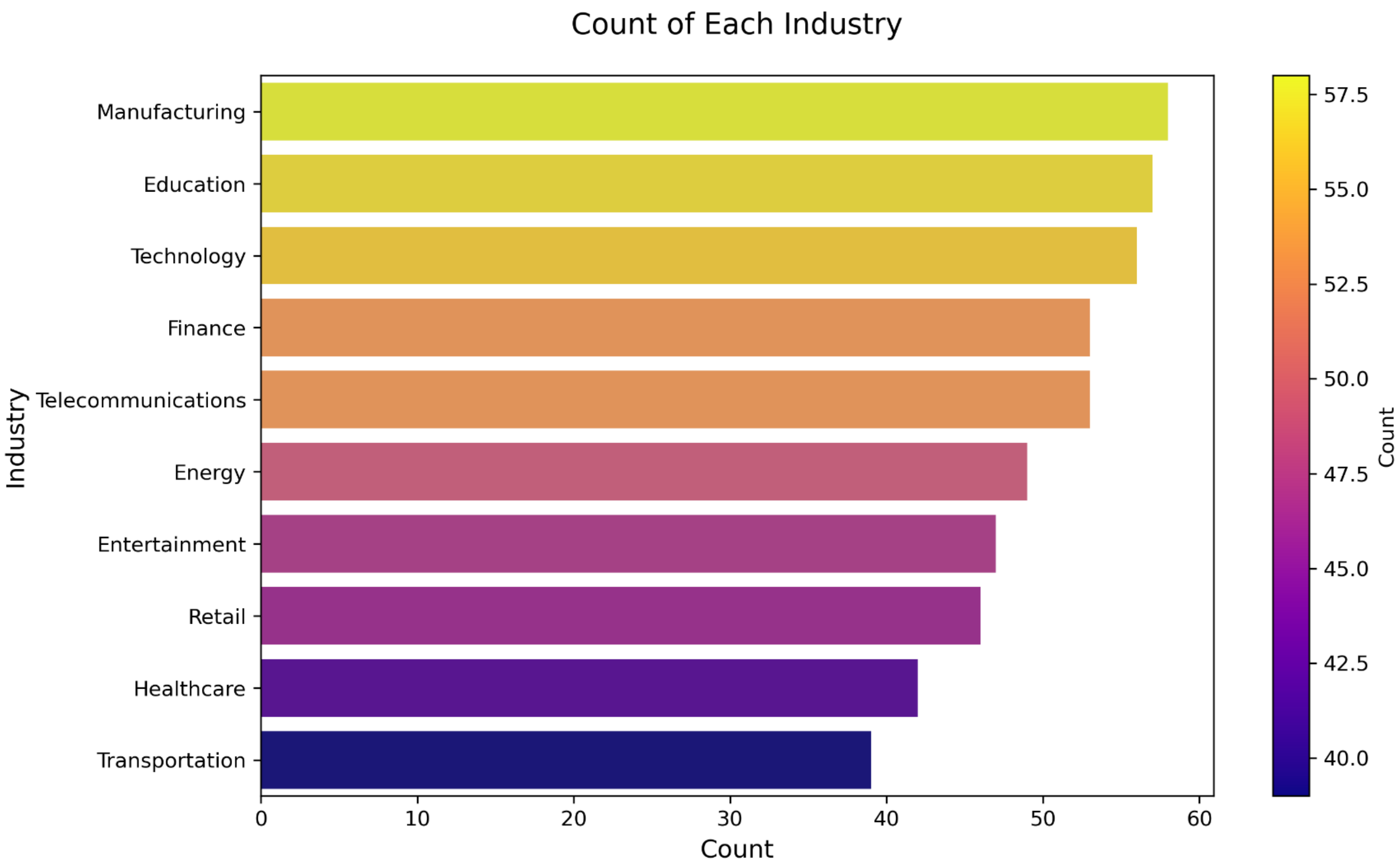
Job Title



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.

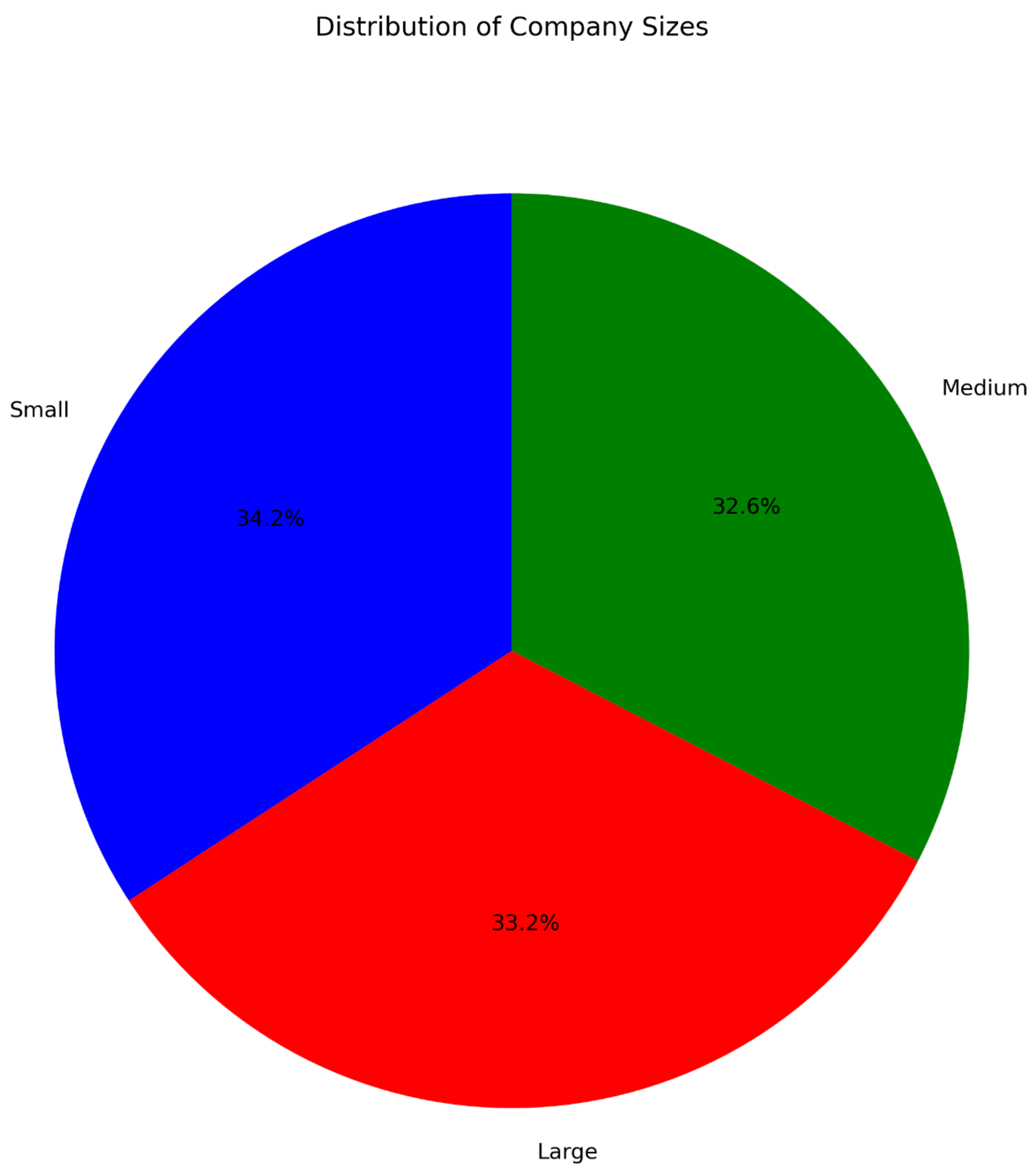
Industry



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.

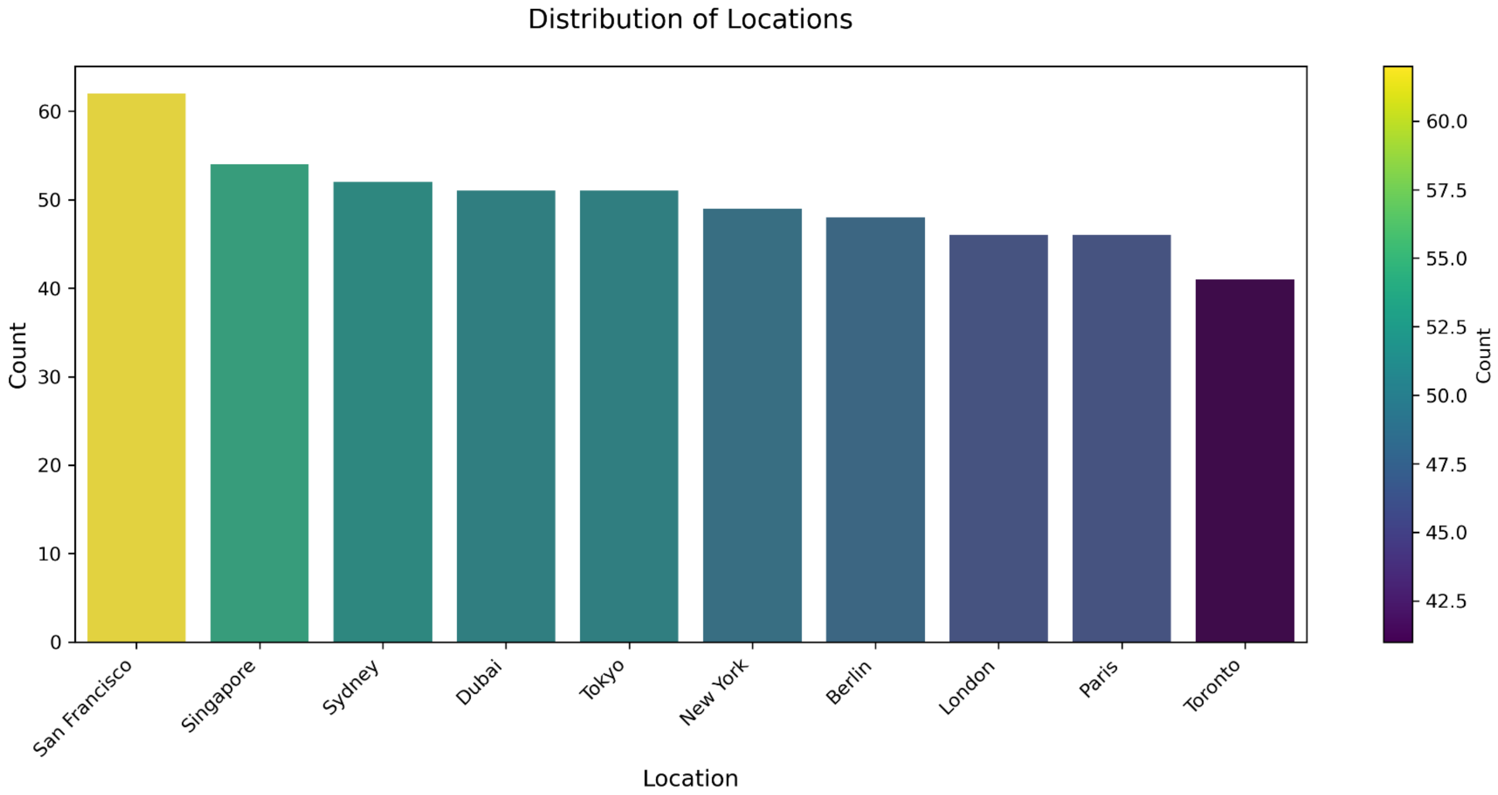
Company Size



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.

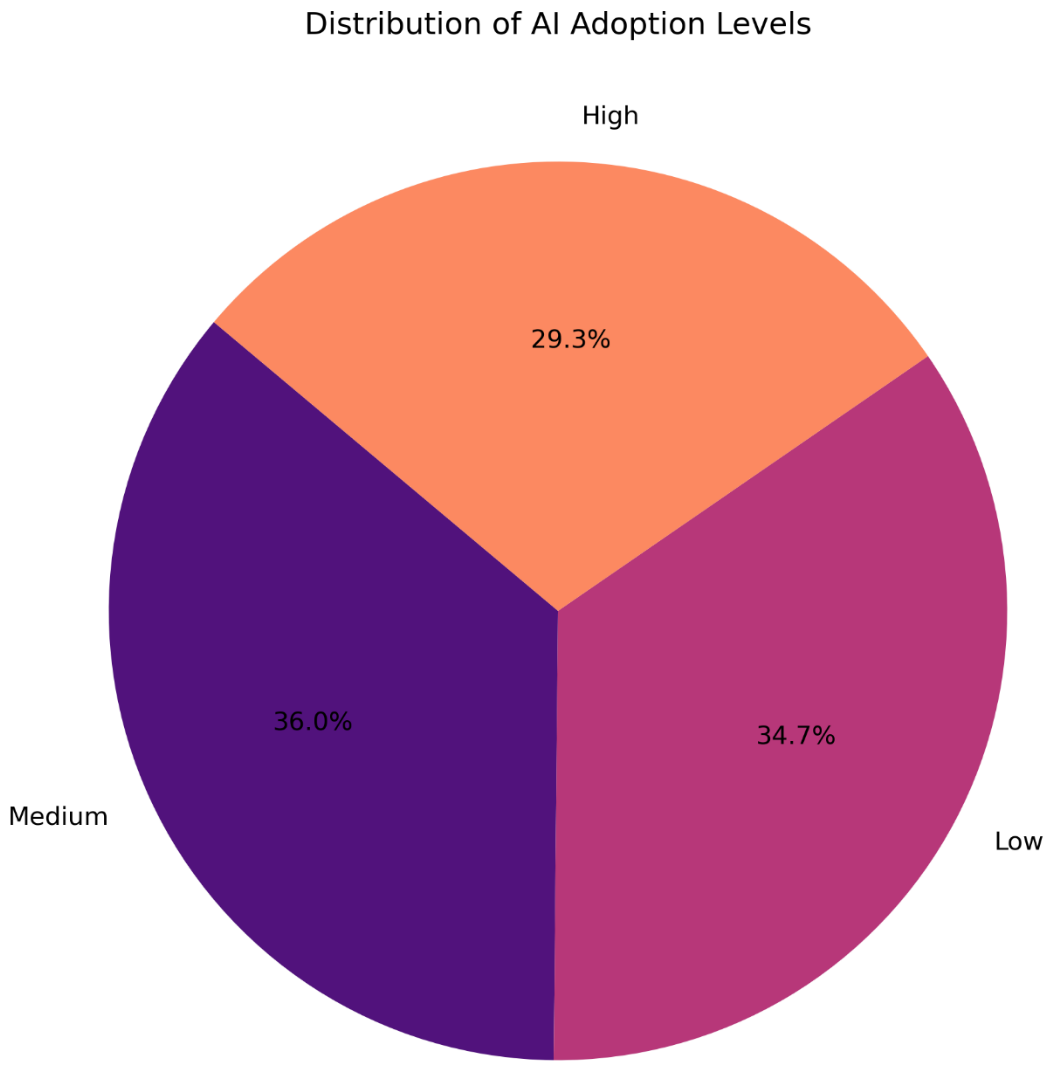
Distribution of location



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.

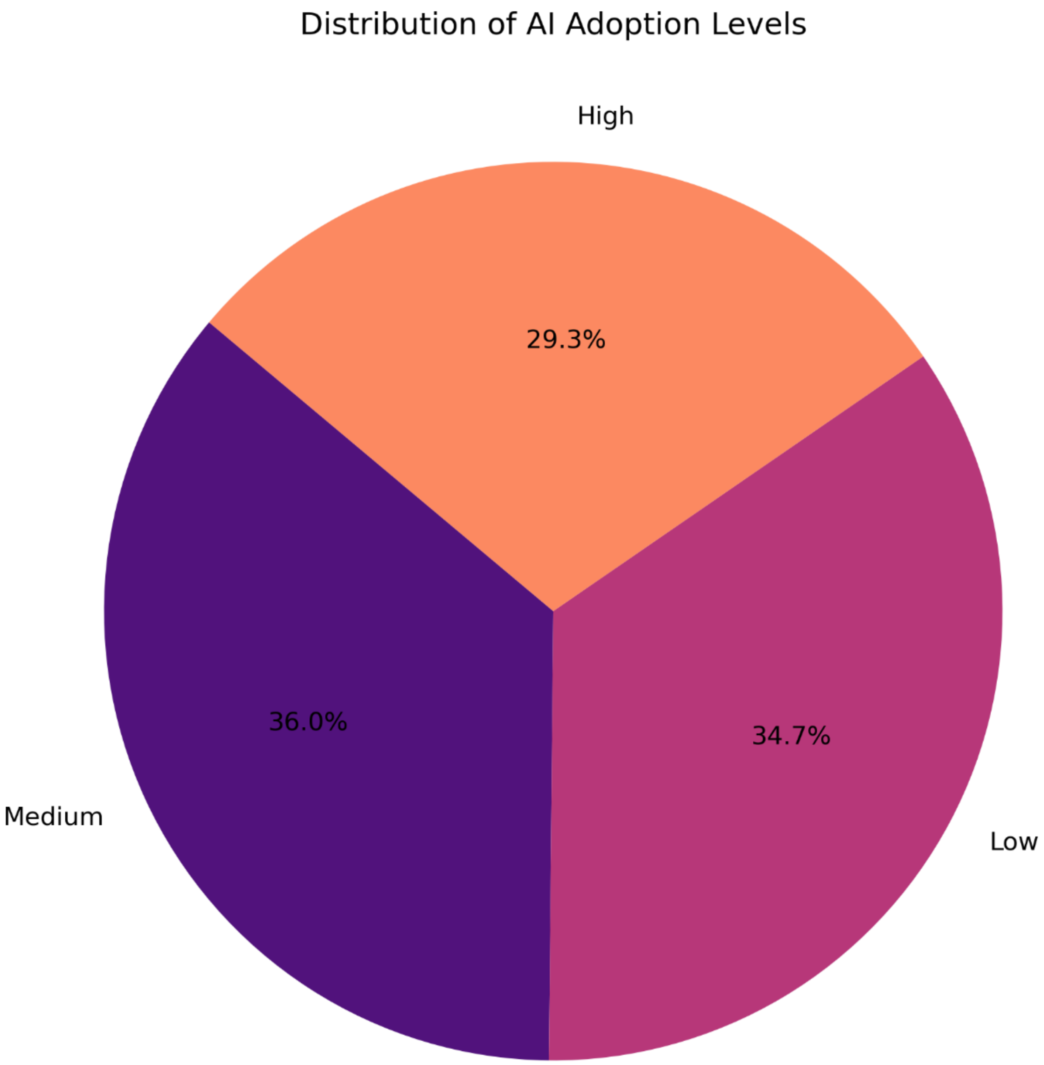
Distribution of location



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.

Distribution of location



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.