Process Book: NextGen Jobs - AI’s Impact on Work

Team Members

- Tailang Cao, u1480633@umail.utah.edu
- Yijun Zhan, u1475897@umail.utah.edu
- Tung hua Chen, u1484206@umail.utah.edu

Table of Contents

1. Introduction

Project Overview:
This project explores the impact of AI on the global job market through a combination of advanced visualizations. By leveraging the "AI-Powered Job Market Insights" dataset, we aim to shed light on how AI adoption influences industries, automation risks, salary trends, skills demand, and job growth projections.

Motivation:
Understanding AI's influence is critical for job seekers, policymakers, and businesses. This project provides data-driven insights to guide workforce decisions in the evolving AI landscape.

2. Project Objectives

The primary goals were:

1. Analyze AI adoption across industries.
2. Assess automation risks for various job roles.
3. Correlate AI adoption with salary trends.
4. Highlight the most in-demand skills in AI-intensive roles.
5. Visualize job growth projections and regional job distribution.

3. Data Processing

3.1 Data Cleaning

Approach:

- Handling Missing Values: We addressed missing values by either filling them with averages (numerical data) or imputing "Unknown" (categorical data).
- Removing Duplicates: Duplicates were removed using Python’s `pandas` library, ensuring data integrity.

- Outlier Detection: We utilized visualization (box plots) to identify and exclude outliers in salary and automation risk.

Key Decisions:

- Retained records with valid salary ranges to ensure meaningful analysis.
- Normalized text fields like job titles and industries using string cleaning techniques.

3.2 Data Transformation

Approach:

- Mapped categorical variables like AI adoption levels into ordinal values for ease of analysis.
- Created new columns:
    - Average salary by industry
    - Automation risk scores derived from job characteristics.
    - Skill frequency scores using text processing on the required skills column.

Key Decisions:

- Grouped industries logically to avoid excessive fragmentation in visualizations.

3.3 Data Aggregation

Approach:

- Aggregated metrics like average salary, automation risk, and skill demand at:
    - Industry level
    - Geographic level
    - AI adoption tiers (e.g., Low, Medium, High).

3.4 Data Filtering

Approach:

- Focused only on jobs from the technology, healthcare, and manufacturing sectors as they dominate AI adoption.
- Excluded irrelevant data, such as job listings lacking sufficient details about AI adoption.

4. Visualization Design

4.1 Iterative Prototyping

Design Process:

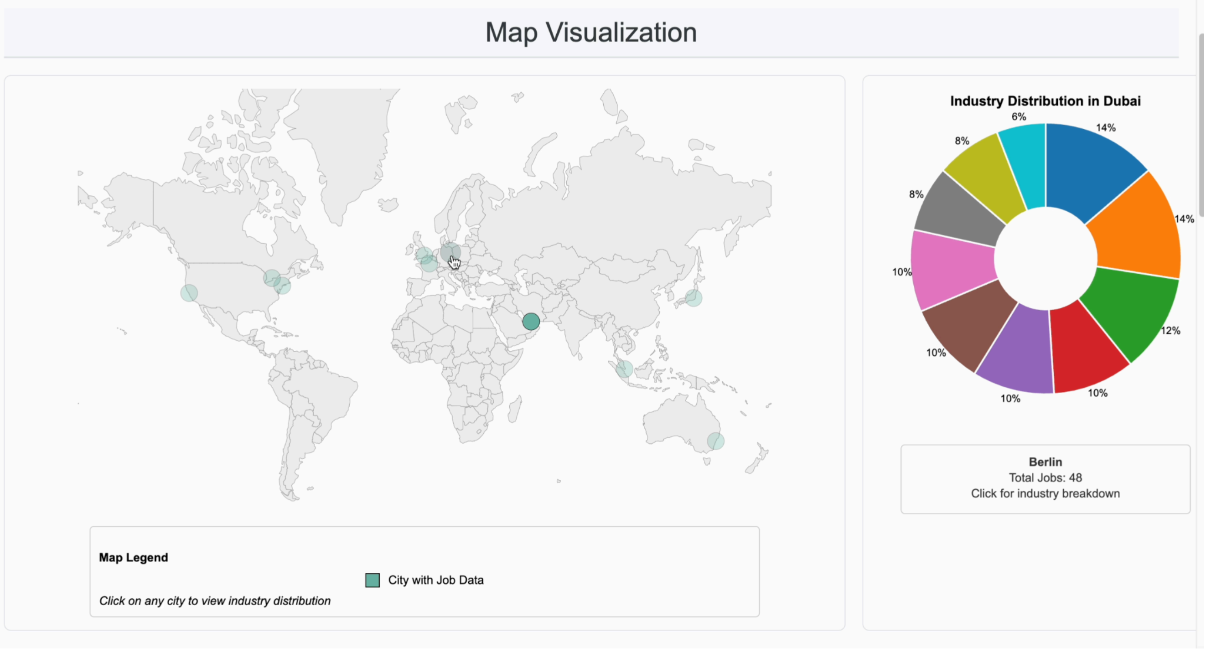
- Brainstormed three prototypes:
  1. A multi-tab dashboard for focused data exploration.
  2. A storytelling format for sequential narrative.
  3. An interactive map for geospatial insights.

Feedback Incorporated:

- Emphasized interactivity in final designs.
- Selected multi-tab dashboard as the core layout for its flexibility.

4.2 Final Visualizations

1. Interactive Map:
   - Purpose: Display job distribution geographically.
   - Tools: GeoJSON for data, D3.js for rendering.
   - Design Choice: Integrated filters for AI adoption, remote-friendliness.

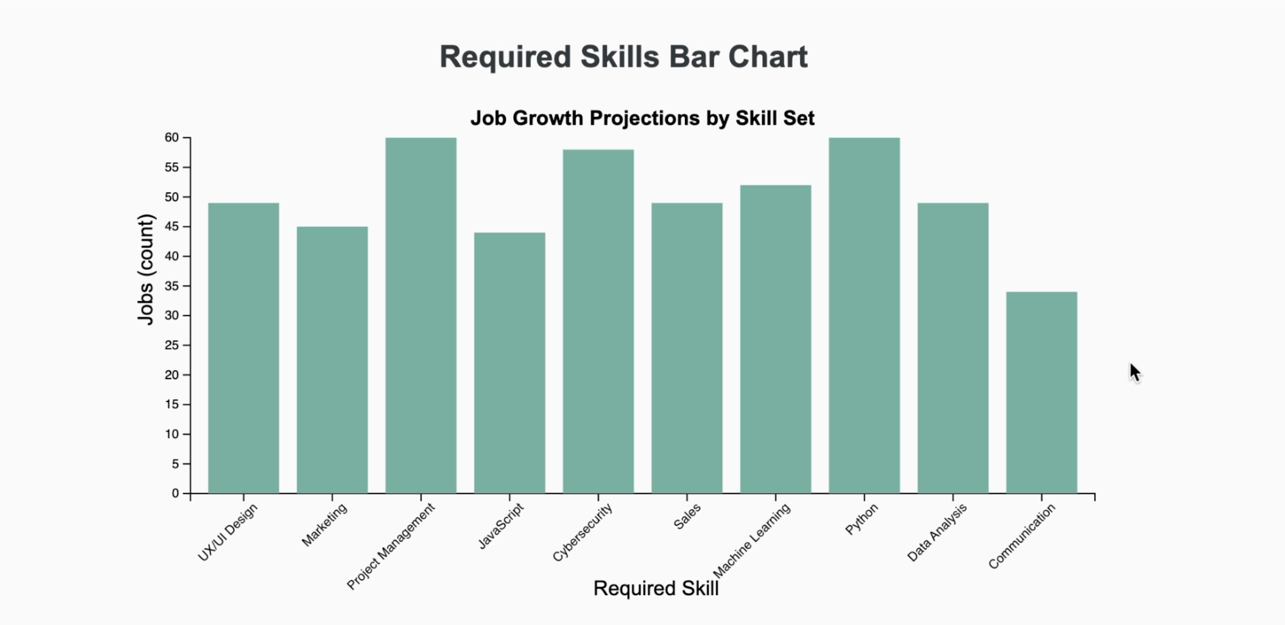


Berlin
Total Jobs: 48
Click for industry breakdown

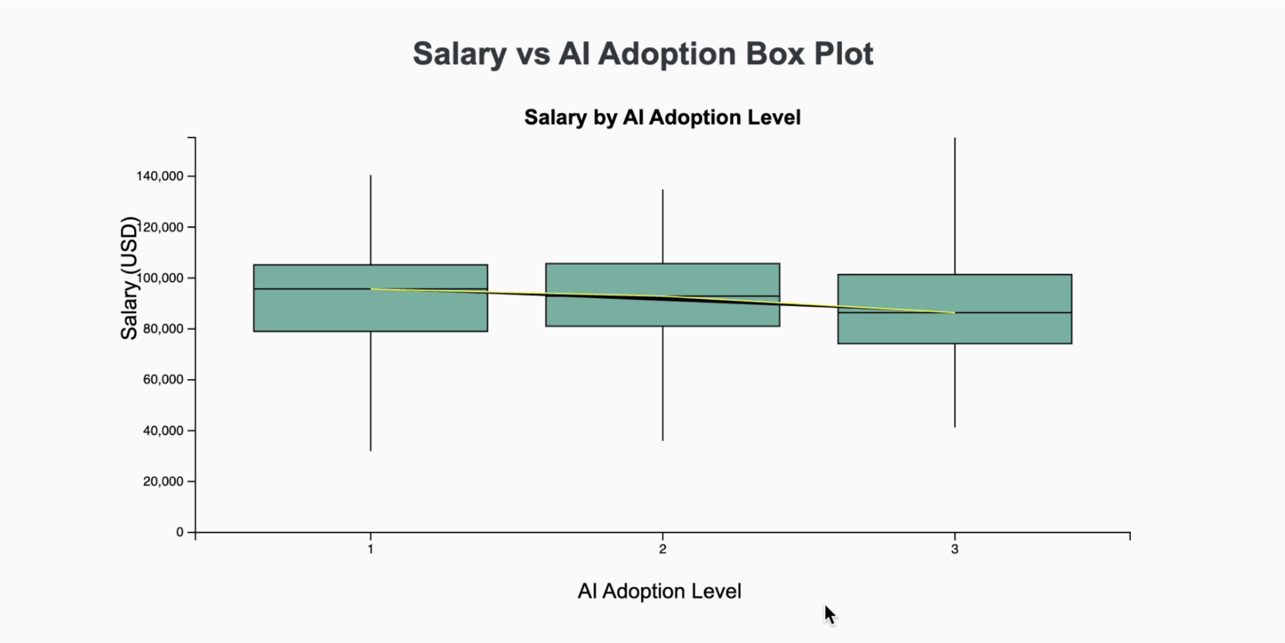
Click on any city to view industry distribution

2. Salary Distribution Box Plots:
   - Purpose: Compare salary ranges across AI adoption tiers.

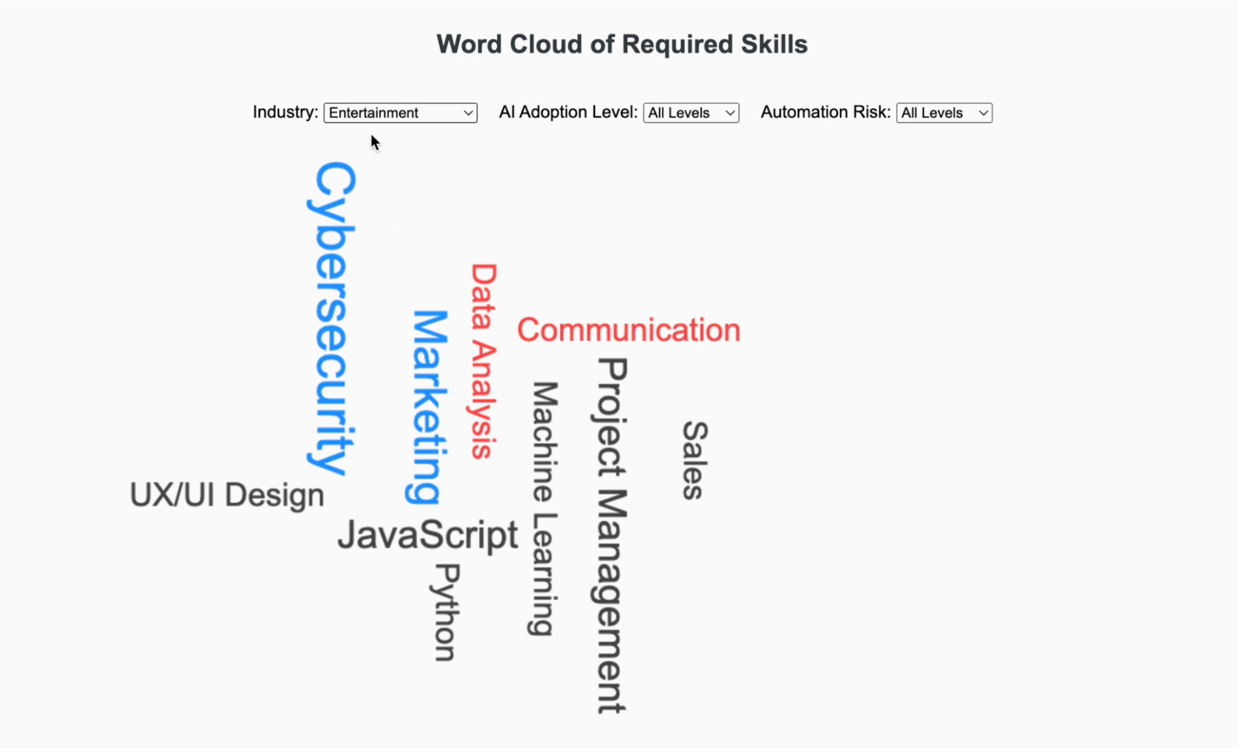
- Tools: Matplotlib (prototype) and D3.js (final).
- Design Choice: Included filters by industry and job category.



3. Skills Demand Word Cloud:
    - Purpose: Highlight skills with size proportional to demand.
    - Tools: `wordcloud` library (Python) for initial processing; D3.js for integration.



4. Job Growth Projections box Graph:
    - Purpose: Show job growth trends over time.
    - Tools: d3.js.
    - Design Choice: Color-coded lines for growth, stability, and decline.



---

5. Implementation

5.1 Tools and Frameworks

- Data Processing: Python (pandas, numpy, matplotlib).
- Visualization Development: D3.js
- Web Framework: HTML, CSS, JavaScript.
- Version Control: GitHub for collaboration and version tracking.

5.2 Workflow

1. Data Preparation: Cleaned and processed the dataset.
2. Prototyping: Created static visualizations for validation.
3. Frontend Integration: Embedded interactive visualizations into the dashboard.
4. Testing: Ensured responsiveness across devices.

5.3 Challenges and Solutions

Here are some challenges we faced while we progressed.

Challenge 1: Aggregating and Transforming Complex Data

- Problem: Some required insights (e.g., job demand trends or average salaries by industry) were not readily available in the raw dataset and needed to be derived from multiple columns.
- Solution:
  - Used Python scripts to aggregate data by grouping it based on specific features like job type, industry, and AI adoption levels.
  - Created new derived metrics, such as average salary per role and AI adoption ratios, to enrich the dataset.
  - Documented all transformations to maintain reproducibility and transparency.

Challenge 2: Overcoming Visualization Design Constraints

- Problem: Ensuring the chosen visualization methods effectively conveyed the required insights while avoiding information overload in the dashboard.
- Solution:
  - Iterated through prototypes, testing designs like heatmaps, bubble charts, and word clouds to balance simplicity with informativeness.
  - Gathered feedback from peers and instructors on prototype usability and adjusted designs accordingly.
  - Prioritized clarity by limiting the number of simultaneous visualizations in each dashboard tab and providing dynamic filters for user-specific customization.

Challenge 3: Performance and Responsiveness

- Problem: The large size of the dataset and the interactivity required for filtering and tooltips slowed down the dashboard, especially on mobile devices.
- Solution:
  - Optimized data loading by pre-aggregating and down-sampling data where high granularity was unnecessary.
  - Used D3.js's efficient rendering capabilities to handle large datasets while maintaining performance.
  - Implemented responsive design practices, such as dynamic resizing and adaptable layouts, to ensure compatibility with different screen sizes.

Challenge 4: Integrating Geospatial Data for Mapping

- Problem: Aligning GeoJSON data with job location information in the dataset was challenging due to discrepancies in location naming conventions.
- Solution:
  - Standardized location names in the dataset to match GeoJSON properties by creating a mapping dictionary.
  - Added fallback mechanisms to handle unmatched locations gracefully by marking them with default or "unknown" values.

Challenge 5: Random Colors in Word Cloud

- Problem: Initially, the word cloud used random colors for all skills, making it difficult to visually distinguish which skills were in higher or lower demand. Users couldn't easily identify key insights from the visualization.
- Solution:
  - Implemented a color-coding scheme to represent skill frequency rankings.
  - The two most frequently required skills were highlighted in blue, while the two least required skills were marked in red.
  - This change improved clarity, allowing users to quickly identify important patterns and prioritize their focus on key skills.

Challenge 6: Usability Testing

- Problem: Ensuring that the dashboard was intuitive for non-technical users while still offering advanced filtering options for deeper insights.
- Solution:
  - Conducted usability testing sessions with peers to identify pain points in navigation and interaction.
  - Simplified user interface elements by adding tooltips, legends, and contextual help popups.
  - Integrated feedback from these sessions into iterative design updates, prioritizing ease of use and clarity.

---

6. Key Findings

- AI Adoption Trends: High adoption in tech and healthcare; manufacturing lags behind.
- Automation Risks: Customer service roles are at the highest risk; creative professions are the least vulnerable.
- Salary Insights: Higher AI adoption correlates with a 20-30% increase in average salary.
- Skills Demand: Machine learning, cloud computing, and data analytics are the most in-demand skills.
- Job Growth: Tech and healthcare sectors show strong growth projections, while manufacturing shows stagnation.

---

7. Future Work

- Optional Features:
  - Sentiment analysis of job descriptions.
  - Predictive modeling for job demand.
  - Animated time series to highlight market changes.
- Usability Improvements:
  - Advanced filtering options (e.g., date ranges).
  - Draggable sliders for dynamic visual updates.

---

8. Conclusion

Impact:

This project successfully bridges the gap between complex job market data and actionable insights. The dashboard provides an intuitive platform for exploring AI's impact on employment, guiding job seekers, businesses, and policymakers alike.

Lessons Learned:

- Collaboration and iterative feedback are vital to creating effective visualizations.
- Striking a balance between data complexity and user-friendliness enhances audience engagement.