Project Report on

AI IMPACT ON JOB MARKET ANALYSIS

In partial fulfillment for the award

Of

Professional Certification in Data Analysis and Visualization

Year 2024-2025

Submitted By

SHAMNA K

Tools & Technologies Used:

Python, R Programming, Tableau

Duration:

15/07/2025 – 30/07/2025

Submission Date:

31/07/2025

G - TEC CENTRE OF EXCELLENCE PERINTHALMANNA

ABSTRACT

This project investigates how artificial intelligence (AI) is influencing the modern job market by analysing a comprehensive dataset of 30,000 job roles across industries, education levels, and countries. The objective is to identify roles most at risk of automation, understand sectoral differences, and explore whether factors such as salary, education, gender diversity, and remote adaptability contribute to vulnerability. Data was analysed using Python and R for statistical insight, while Tableau was used to create interactive dashboards. The findings reveal that over 30% of jobs are at high risk of automation, with industries like manufacturing and healthcare facing the greatest transformation. This project highlights critical trends that can inform workforce planning and reskilling strategies.

INTRODUCTION

Artificial Intelligence (AI) is rapidly transforming the global job market, reshaping demand for skills, redefining roles, and influencing employment trends across nearly every sector. While AI brings promises of increased efficiency and innovation, it also raises concerns around job displacement and the automation of routine tasks. In this context, understanding how AI is impacting various industries, job roles, and education requirements is critical for policymakers, job seekers, and organizations preparing for the future of work.

This project, titled *"AI Impact on the Job Market,"* explores how AI is influencing employment across different dimensions including automation risk, AI impact levels, projected job openings, salaries, experience requirements, and gender diversity. Using a dataset comprising 30,000 synthetic job records modelled on real-world trends and labour reports from reputable sources such as the World Economic Forum and McKinsey Global Institute, the project aims to deliver meaningful insights into which job profiles are at greater risk and which may see growth due to AI advancements.

The analysis was conducted using Python for data cleaning, feature engineering, and exploratory data analysis; R programming for statistical hypothesis testing; and Tableau for creating interactive dashboards and visualizations. This multi-tool approach allowed for a comprehensive understanding of the data, revealing patterns across industries, regions, education levels, and career stages. Through this study, we seek to uncover how factors such as salary, experience, and education correlate with automation risk and job stability in the age of AI, providing stakeholders with a clearer view of emerging trends and future workforce implications.

LITERATURE REVIEW

1. **Autor et al. (2003)** explored how computerization affects labour demand, noting that while routine jobs are increasingly automated, there is growing demand for non-routine cognitive and interpersonal skills. This introduced the concept of "task polarization," which has continued to shape workforce development and education policy over the past two decades.
2. **Autor and Dorn (2013)** provided further empirical evidence of labour market polarization in the United States. Their research highlighted how advances in computer technology were displacing middle-skill, routine-based jobs—particularly in sectors such as manufacturing and administration—while both high-skill analytical and low-skill service jobs were expanding in response to changes in labour demand.
3. **Frey and Osborne (2017)** estimated that 47% of total US employment was at risk of automation. Using a task-based modelling approach, they evaluated over 700 occupations and found that roles requiring perception, creativity, and social intelligence were more resistant to automation. Their study became foundational in automation risk modelling and has influenced both academic and policy discussions globally.
4. **World Economic Forum (2020)**, in its *Future of Jobs* report, suggested that while automation may displace up to 85 million jobs by 2025, it could simultaneously create 97 million new roles that are better aligned with the evolving division of labour between humans and machines. The report emphasized the urgency of large-scale reskilling and upskilling initiatives across economies.
5. **McKinsey Global Institute (2021)** projected that by 2030, up to 375 million workers globally might need to transition into new occupations due to automation. The study further noted that the impact of AI adoption will vary depending on industry structures and regional economic conditions, highlighting the need for targeted labour policies.
6. **Chui, Manyika, and Miremadi (2021)** conducted a task-level analysis across multiple industries and concluded that while very few jobs are fully automatable, roughly 60% of occupations contain at least 30% of tasks that could be automated using currently available technologies. Their findings provided a more nuanced understanding of partial automation risk.
7. **OECD (2022)** offered a global comparative perspective, observing that automation risk is significantly higher in economies with a greater share of low-skilled jobs. The report

also examined gender disparities, noting that female-dominated roles may face disproportionate automation exposure due to the nature of tasks involved.

8. **ILO (2023)** analysed how automation and AI could shape employment in developing countries, warning that without supportive policies, these technologies could intensify job insecurity and income inequality. The report advocated for coordinated responses including social protection systems, inclusive digital access, and workforce reskilling to promote equitable adaptation.
9. **Brynjolfsson and McAfee (2024)** revisited their earlier work on the "second machine age" to examine how generative AI is now affecting cognitive and creative occupations. They argue that while AI is not eliminating all jobs, it is changing the nature of work, and the emerging challenge lies in aligning the existing workforce's skills with newly evolving roles.

RESEARCH GAP

Despite extensive research on the impact of artificial intelligence and automation on the labour market, much of the existing literature remains either high-level or focused on specific regions or industries, often lacking a multidimensional view of the workforce. Studies by Frey and Osborne (2017), McKinsey Global Institute (2021), and the World Economic Forum (2020) have projected widespread job disruption but seldom integrate variables such as education level, gender diversity, remote work feasibility, and experience level within a unified analysis. Additionally, most analyses rely on task-level models or macroeconomic estimates and do not provide granular, data-driven insights that can inform practical decision-making. This project addresses that gap by incorporating feature engineering, hypothesis testing, and interactive visualizations to explore how AI influences job risk, salary trends, and future demand across industries and demographics, offering a more comprehensive and accessible approach to understanding AI's evolving role in the labour market.

DATA COLLECTION & PRE-PROCESSING

Data Source and Collection Methods

The dataset used in this project is ai_job_trends_dataset.csv sourced from Kaggle. It contains 30,000 job records across various industries and countries, with fields capturing details about AI impact, job growth, salary, automation risk, and job characteristics. The dataset simulates real-world labour trends based on global research from sources such as the World Economic Forum, OECD, and McKinsey. It was provided in CSV format and imported into Python for analysis. The data covers multiple dimensions including AI impact level (Low, Moderate, High), median salary, job status (increasing or decreasing), education level, gender diversity, and projected job openings for 2024 and 2030.

Data Quality Assessment and Cleaning Procedures

Initial inspection and cleaning were performed in Python using the pandas and numpy libraries. The following steps were taken:

- **Missing values**: Checked and confirmed that no null or missing entries were present in the dataset.
- **Duplicate entries**: Verified and removed any duplicate rows to ensure data integrity.

Feature Engineering and Selection Techniques

Several new features were created to support deeper analysis and testing:

- **Growth_Rate(%)**:
  Calculated as the percentage change between Projected Openings in 2030 and Openings in 2024 to assess demand trends.
- **Experience_Level**:
  Groups years of experience into categories:
  - Entry-Level: 0–3 years
  - Mid-Level: 4–8 years
  - Senior-Level: 9+ years

- **High_Risk_Flag**:
  Jobs with Automation Risk > 75% were labelled as "High Risk" for targeted filtering and analysis.

Columns Selected for Analysis

The following key columns were used in most of the analysis and visualization:

- Job_Title
- Industry
- AI_Impact_Level
- Job_Status
- Median_Salary
- Projected_Openings_2030
- Job_Openings_2024
- Automation_Risk
- Required_Education
- Experience_Required
- Remote_Work_Ratio
- Gender_Diversity

METHODOLOGY

This project uses a multi-tool analytical approach involving Python, R programming, and Tableau to study the impact of Artificial Intelligence (AI) on the job market. The methodology was designed to enable robust data cleaning, exploratory analysis, statistical validation, and intuitive presentation of insights. The combination of tools ensures both technical accuracy and accessibility for decision-makers.

Tools and Technologies Used

- Python was used for initial data exploration, cleaning, transformation, and feature engineering. Libraries such as pandas, numpy, matplotlib,seaborn and plotly supported tasks like missing value handling, trend identification, and outlier detection.
- R programming was utilized for conducting statistical hypothesis testing, confirming whether the trends observed in the exploratory analysis were statistically significant. Key functions used include t.test(), chisq.test(), aov(), var.test() and z.test().
- Tableau served as the platform for dashboard creation and data visualization, allowing stakeholders to interact with the findings using filters and graphical summaries. Its user-friendly design helped communicate complex AI-related job market trends in a visual format.

Exploratory Data Analysis – Python

- Once the dataset was cleaned, exploratory data analysis (EDA) was carried out to reveal patterns and correlations. This step was essential in guiding the design of the Tableau dashboard and the choice of statistical tests in R.

Visualizations:

- **Pie Chart**: Used to find top 10 industries with high_risk jobs.
- **Bar Chart**: Help to find top 10 high_risk job title.
- **Vertical Bar Plot:** To understand how job growth trends differ based on the level of AI impact a job faces and To analyse whether career stage (Entry, Mid, Senior) affects the **risk of automation**.
- **Line Chart**: To understand how future job demand (growth from 2024 to 2030) varies across career stages (Entry, Mid, Senior).

- **Groupby:**

  The groupby() function was used to analyse average automation risk, median salary, and job growth by industry and education level.

Statistical Testing – R Programming

To validate EDA insights, hypothesis testing was conducted using R. This step helped assess whether observed differences were statistically significant.

- **T-test**: Test whether Median Salary differs significantly between High vs Low Gender Diversity groups
- **Chi-square Test**: checks whether there is a significant association between the type of industry and the education level required.
- **ANOVA**: checks if the mean number of job openings (2024) differs significantly across different levels of required education.
- **Z-test**: Compare average Automation Risk (%) significantly different from 50
- **F-test**: Compare high-risk and low-risk job groups have different variance in salary

Dashboarding and Visualization – Tableau

After completing the data analysis, Tableau was used to create an **interactive dashboard** for exploring AI job market impacts. The dashboard was designed for usability, with clear layout, colour coding, and responsive filters.

Key Features:

Charts Included:

- **Projected Openings Over Time**: A line chart comparing job openings in 2024 and projected values in 2030.
- **Top 10 Declining Jobs**: A horizontal bar chart showing job roles with the steepest decrease in demand.
- **Job Status by Industry**: A grouped bar chart highlighting the count of jobs growing vs. shrinking by sector.

- **Automation Risk by Experience**: A bubble plot showing how automation risk varies by career stage.
- **Education Level Distribution**: A donut chart presenting the breakdown of required education across job types.

KPI Cards:

- **Average Salary**: Displays the average median salary across the dataset.
- **% High-Risk Jobs**: Indicates the proportion of jobs with automation risk above the threshold.
- **Total Jobs**: Shows the total number of job records analysed.

Interactive Filters:

- **Industry Filter**: Enables users to drill down into specific sectors.
- **AI Impact Level Filter**: Allows analysis based on AI exposure categories (Low, Moderate, High).

RESULTS AND ANALYSIS

This section presents the key findings of the project, focusing on how different job roles, industries, and educational backgrounds are affected by artificial intelligence (AI) and automation.

Python-Based Results

- Roles such as Fast Food Manager *and* Electrical Engineer *show near*-100% automation risk, making them most vulnerable to AI displacement.
- Several high-risk roles, including Mechanical Engineer and Oncologist, are also experiencing steep declines in job openings.
- Entertainment, healthcare, and manufacturing industries have the highest average automation risk and the largest number of high-risk jobs.
- Higher salaries do not protect jobs from automation, as high-risk roles earn nearly as much as lower-risk ones.
- All education levels— including PhDs—face significant automation risk, indicating no qualification guarantees immunity.
- Jobs like Sports Coach and Jewellery Designer show explosive growth potential, while roles such as Publishing Copy Editor are nearly vanishing.
- Automation risk remains consistent across entry-level to senior-level jobs, suggesting experience does not shield against AI disruption.
- There is no strong relationship between remote work availability and automation risk, making telework a neutral factor.
- The most affected sectors include entertainment, healthcare, and IT, each with about 12–13% of all high-risk jobs.
- Countries like the UK, USA, and Australia lead in the number of high-risk jobs, indicating global exposure to AI automation.
- Gender diversity does not significantly influence automation risk, with risk levels nearly equal across gender-balanced and imbalanced roles.

R-Based Statistical Results

Statistical testing was used to validate the patterns observed in the Python analysis:

- T-test shows a statistically significant difference in median salary between high and low gender diversity job groups, with higher diversity roles earning $821.79 more on average.
- Chi-square Test Revealed that the level of education required for a job varies significantly by industry, suggesting industry-specific educational expectations.
- ANOVA Showed there is a statistically significant difference in the average number of job openings (2024) across different levels of required education.
- Z-test: Showed On average, jobs in the dataset face a moderate automation risk, with no significant deviation from the expected 50% risk level.
- F-test shows there is no statistically significant difference in salary variance between high-risk and low-risk job groups.

Tableau-Based Results

An interactive Tableau dashboard was created to visualize and communicate the findings:

- While some job roles are projected to grow steadily by 2030, others show a notable decline in future openings, reflecting AI-driven disruption.
- The largest declines are seen in specific roles, with some jobs expected to decrease by over 500 positions, highlighting areas most susceptible to AI replacement.
- Jobs are evenly spread across education levels, with Bachelor's and Master's degrees being slightly more dominant, suggesting AI impact spans all qualification tiers.
- Surprisingly, Senior-Level roles have the highest automation risk (50.28%), indicating that even experience doesn't guarantee protection from AI automation.
- Both growing and shrinking roles are evenly distributed across industries, implying that AI is transforming rather than uniformly eliminating industry-specific jobs.

CONCLUSION

This study presents an in-depth exploration of how artificial intelligence (AI) is transforming the dynamics of the global job market. Through rigorous analysis of job characteristics, risk assessments, and employment projections, it becomes evident that the risk of automation is no longer confined to routine or low-skilled roles. In fact, the findings show that even senior-level positions and jobs requiring advanced education, such as those with Master's or PhDs, are increasingly vulnerable to automation technologies. This signals a paradigm shift—experience and education alone no longer serve as safeguards against AI disruption.

Furthermore, the analysis of projected job growth and decline by 2030 reveals a dual reality: while certain roles are set to experience rapid expansion, others—particularly those with high automation risk—are expected to shrink dramatically. Some occupations show a predicted decrease of hundreds of job openings, underscoring how automation could lead to large-scale displacement in specific fields. Notably, these trends are not isolated to one sector; industries ranging from manufacturing and finance to healthcare and entertainment all exhibit varying degrees of exposure to AI-driven changes.

The study also highlights that automation risk is spread across different education levels and job statuses, challenging the traditional assumption that technological advancements primarily affect low-skilled labour. Even jobs with high median salaries and senior experience levels are at risk, indicating that AI is reshaping not only what jobs exist, but also who is most affected by these shifts.

These insights collectively emphasize the need for a strategic and forward-thinking response. Governments, industries, and educational institutions must collaborate to implement policies that encourage lifelong learning, re-skilling, and agile workforce development. Investing in human adaptability—through education reform, digital literacy programs, and targeted training—will be critical to equipping workers for the future of work.

In essence, the impact of AI on employment is not merely about replacement; it's about transformation. As technology continues to evolve, the nature of work will increasingly demand adaptability, creativity, and continuous learning. The future workforce must be prepared not just to survive—but to thrive—in an economy fundamentally reshaped by intelligent automation.

FUTURE WORK

- Use forecasting models to better predict job trends.
- Analyse skill gaps in high-risk job categories.
- Simulate government policies (like reskilling programs) to test their impact.
- Compare AI's impact across different countries and regions.
- Use real-time job data for more accurate insights.
- Study how AI changes tasks instead of just replacing jobs.
- Explore social and ethical issues like fairness, job access, and worker well-being.

REFERENCES

- Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. The Quarterly Journal of Economics, 118(4), 1279–1333.
- Autor, D., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labour market. American Economic Review, 103(5), 1553–1597.
- Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? Technological Forecasting and Social Change, 114, 254–280.
- World Economic Forum. (2020). The Future of Jobs Report 2020
- McKinsey Global Institute. (2021). The future of work after COVID-19.
- Chui, M., Manyika, J., & Miremadi, M. (2021). A future that works: Automation, employment, and productivity. McKinsey Global Institute.
- OECD. (2022). Automation, skills use and training. OECD Publishing.
- International Labour Organization (ILO). (2023). AI and the future of work: Global challenges and responses. Geneva: ILO Publications.
- Brynjolfsson, E., & McAfee, A. (2024). Generative AI and the second machine age revisited. MIT Initiative on the Digital Economy.

SUPPORTING FILES

Python

```
import pandas as pd
import numpy as np
import warnings
warnings.filterwarnings("ignore")
```

```
df=pd.read_csv("./ai_job_trends_dataset.csv")
df.head()
```

```
df.info()
```

```
df.describe()
```

```
df.shape
```

```
df.columns
```

```
df.isnull()
```

```
df.isnull().sum()
```

```
df.duplicated().sum()
```

Identify Jobs at High Risk of AI Automation ¶

(Discover which roles are most exposed to AI-driven automation.)

```
###Filter jobs with automation risk greater than 70%
```

```
high_risk_jobs = df[df['Automation Risk (%)'] > 70]
high_risk_jobs
```

```
# show top 10 jobs with highest automation risk
```

```
high_risk_jobs[['Job Title', 'Industry', 'Automation Risk (%)']]\
    .sort_values(by='Automation Risk (%)', ascending=False).head(10)
```

```
# Insight: These are the job titles most likely to be replaced or transformed by AI.
```

Growth vs. Risk Analysis

(whether high-risk jobs are also expected to decline in availability.)

```
# calculate job growth between 2024 to 2030
```

```
df['Growth'] = df['Projected Openings (2030)'] - df['Job Openings (2024)']
df['Growth']
```

```
# Display 10 most declining high-risk jobs
```

```
growth_risk = df[df['Automation Risk (%)'] > 70][['Job Title', 'Growth', 'Automation Risk (%)']]
growth_risk.sort_values(by='Growth').head(10)
```

Industry-Wise AI Risk Impact

Understand which industries have the highest average risk and concentration of high-risk jobs.

```
# Average automation risk per industry
```

```
industry_risk = df.groupby('Industry')['Automation Risk (%)'].mean().sort_values(ascending=False)
industry_risk
```

```
# Count of high-risk jobs per industry
```

```
high_risk_industry = df[df['Automation Risk (%)'] > 70].groupby('Industry')
['Job Title'].count().sort_values(ascending=False)
high_risk_industry
```

```
# Merge both metrics into a single DataFrame
```

```
industry_impact = pd.DataFrame({
    'Avg Automation Risk (%)': industry_risk,
    'High-Risk Job Count': high_risk_industry
}).fillna(0).sort_values(by='High-Risk Job Count', ascending=False)
industry_impact
```

```
# Insight: Some industries may be more exposed to automation (e.g., Manufacturing, Retail, Transport).
```

Salary Comparison _High Risk vs Low Risk

(explore how automation risk correlates with income)

```
#create risk category column

df["Risk Category"] = df['Automation Risk (%)'].apply(lambda x: "High Risk"
                   if x > 70 else "Low/Moderate Risk")
```

```
#  compare average salary by risk group
salary_risk_comparison= df.groupby("Risk Category")["Median Salary (USD)"].mean().sort_values()
print(salary_risk_comparison)
```

```
# AI automation risk affects jobs across all pay levels, showing salary is not a safeguard against automation.
```

Education Level vs. Automation Risk

#(Investigate whether education level affects the likelihood of AI automation.)

```
# Average automation risk by education level
```

```
education_risk = df.groupby('Required Education')['Automation Risk (%)'].mean().sort_values()
education_risk
```

```
# Count of high-risk jobs per education level
```

```
high_risk_edu = df[df['Automation Risk (%)'] > 70].groupby('Required Education')
['Job Title'].count().sort_values(ascending=False)
high_risk_edu
```

```
# Merge into one table
```

```
education_impact = pd.DataFrame({
    'Avg Automation Risk (%)': education_risk,
    'High-Risk Job Count': high_risk_edu
}).fillna(0).sort_values(by='Avg Automation Risk (%)')
education_impact
```

```
# Higher education may offer more protection from automation, but it's not absolute.
```

Projected Growth Rate Calculation

```
df['Projected Growth Rate (%)'] = ((df['Projected Openings (2030)'] - df['Job Openings (2024)'])
                                  / df['Job Openings (2024)']) * 100
```

```
# If the result is positive, the job is growing

# If the result is negative, the job is shrinking
```

```
# Top 10 Growing Jobs
df[['Job Title', 'Projected Growth Rate (%)']].sort_values(by='Projected Growth Rate (%)', ascending=False).head(10)
```

```
#Top 10 declining jobs
df[['Job Title', 'Projected Growth Rate (%)']].sort_values(by='Projected Growth Rate (%)').head(10)
```

```
# A negative growth rate means the number of job openings in 2030 is expected to be lower than in 2024.
# A rate of -98% suggests the job is almost disappearing by 2030.
```

Experience Level

```
def categorize_experience(x):
    if pd.isna(x):
        return 'Unknown'
    elif x <= 3:
        return 'Entry-Level'
    elif 4 <= x <= 8:
        return 'Mid-Level'
    else:
        return 'Senior-Level'

df['Experience Level'] = df['Experience Required (Years)'].apply(categorize_experience)
print(df[['Experience Required (Years)', 'Experience Level']].head())
```

```
# Average automation risk by experience level:
```

```
df.groupby('Experience Level')['Automation Risk (%)'].mean()
```

```
# automation risk does not decrease with experience. Even senior professionals
# may be affected by AI and automation, depending on the nature of their jobs.
```

Remote Work vs Automation Risk

(Check if jobs with higher automation risk are more or less likely to allow remote work.)

```
remote_risk = df.groupby('Risk Category')['Remote Work Ratio (%)'].mean().sort_values(ascending=False)
print(remote_risk)
```

```
# AI-vulnerable jobs are not significantly more or less remote than safer jobs.

# Remote work doesn't appear to influence automation risk in this dataset.
```

High-Risk Job Insights

- a) Top 10 Industries

```
# e jobs with Automation Risk (%) > 75 as "High Risk":
df['High Risk Flag'] = df['Automation Risk (%)'].apply(lambda x: 'High Risk' if x > 75 else 'Not High Risk')
print(df['High Risk Flag'])
```

```
high_risk_jobs = df[df['High Risk Flag'] == 'High Risk']
```

```
top_10_industries = high_risk_jobs['Industry'].value_counts().head(10)
print(top_10_industries)
```

```
import pandas as pd
import plotly.express as px

# Step 1: Filter high-risk jobs
high_risk_jobs = df[df['High Risk Flag'] == 'High Risk']

# Step 2: Get top 10 industries by count, convert to DataFrame
top_10_industries_df = high_risk_jobs['Industry'].value_counts().head(10).reset_index()

# Step 3: Rename columns correctly
top_10_industries_df.columns = ['Industry', 'count']  # 'Industry' is now a column

# Step 4: Create pie chart
fig = px.pie(
    top_10_industries_df,
    names='Industry',
    values='count',
    title='Top 10 Industries with High-Risk Jobs (Pie Chart)',
    template='plotly_white'
)

fig.show()
```

```
# b) Required Education Distribution
```

```
print(df[df['High Risk Flag'] == 'High Risk']['Required Education'].value_counts())
```

```
# c) Gender Diversity Descriptive Stats
```

```
print(df[df['High Risk Flag'] == 'High Risk']['Gender Diversity (%)'].describe())
```

Gender Diversity vs Automation Risk

(Explore if automation risk impacts jobs with higher or lower gender diversity.)

```
gender_risk = df.groupby('Risk Category')['Gender Diversity (%)'].mean().sort_values(ascending=False)
print(gender_risk)
```

Job Status vs AI Impact

(See whether jobs with high automation risk are increasing or decreasing in demand.)

```
job_status_risk = df.groupby('Job Status')['Automation Risk (%)'].mean().sort_values(ascending=False)
print(job_status_risk)
```

Experience Level vs Automation Risk

Investigate if AI risk is higher for entry-level or experienced roles.

```
experience_risk = df.groupby('Experience Required (Years)')['Automation Risk (%)'].mean().sort_values()
print(experience_risk)
```

```
# Automation risk remains consistent across experience levels, indicating AI impacts entry-level and senior roles alike,
# with only slight variation.
```

Growth Forecast by Industry for High-Risk Jobs

(Check which industries have growth despite being high-risk.)

```
high_risk_growth = df[df['Automation Risk (%)'] > 70].groupby('Industry')['Growth'].mean().sort_values(ascending=False)
print(high_risk_growth)
```

```
# High-risk jobs in healthcare and manufacturing are growing due to AI-driven transformation, while transportation
# and entertainment see declines as automation replaces tasks.
```

Job Titles with High Risk but High Growth

Identify jobs that are risky but still growing (possible transformation roles).

```
high_risk_high_growth = df[(df['Automation Risk (%)'] > 70) & (df['Growth'] > 0)]
print(high_risk_high_growth[['Job Title', 'Industry', 'Growth', 'Automation Risk (%)']].head(10))
```

Bar Chart – Top 10 High-Risk Job Titles

```
import plotly.express as px

# Filter to only high-risk jobs
high_risk_jobs = df[df['High Risk Flag'] == 'High Risk']

# Get top 10 most common high-risk job titles
top_jobs = high_risk_jobs['Job Title'].value_counts().head(10).reset_index()
top_jobs.columns = ['Job Title', 'Count']

# Create bar chart
fig = px.bar(
    top_jobs,
    x='Count',
    y='Job Title',
    orientation='h',
    color='Job Title',
    text='Count',
    title='Top 10 High-Risk Job Titles'
)

fig.update_layout(
    yaxis=dict(autorange='reversed'),
    plot_bgcolor='white',
    showlegend=False
)

fig.show()
```

Bar Chart

```
import seaborn as sns
import matplotlib.pyplot as plt

sns.set(style="whitegrid")
fig, axs = plt.subplots(2, 1, figsize=(10, 12))  # Changed from 3 to 2 rows

# 1. Projected Growth Rate by AI Impact Level
sns.barplot(
    x=df.groupby('AI Impact Level')['Projected Growth Rate (%)'].mean().index,
    y=df.groupby('AI Impact Level')['Projected Growth Rate (%)'].mean().values,
    ax=axs[0], palette='viridis'
)
axs[0].set_title("Avg. Projected Growth Rate by AI Impact Level")
axs[0].set_ylabel("Projected Growth Rate (%)")

# 2. Automation Risk by Experience Level
sns.barplot(
    x=df.groupby('Experience Level')['Automation Risk (%)'].mean().index,
    y=df.groupby('Experience Level')['Automation Risk (%)'].mean().values,
    ax=axs[1], palette='Set2'
)
axs[1].set_title("Avg. Automation Risk by Experience Level")
axs[1].set_ylabel("Automation Risk (%)")

plt.tight_layout()
plt.show()
```

Chart: Average Projected Growth Rate by Experience Level

To understand how future job demand (growth from 2024 to 2030) varies across career stages (Entry, Mid, Senior).

```
import plotly.express as px

# Group by Experience Level
growth_by_exp = df.groupby('Experience Level')['Projected Growth Rate (%)'].mean().reset_index()

# Line plot
fig = px.line(
    growth_by_exp,
    x='Experience Level',
    y='Projected Growth Rate (%)',
    markers=True,
    title='Projected Job Growth Rate by Experience Level',
    template='plotly_white'
)
fig.update_traces(line_color='royalblue', marker=dict(size=10))
fig.show()
```

R Programming

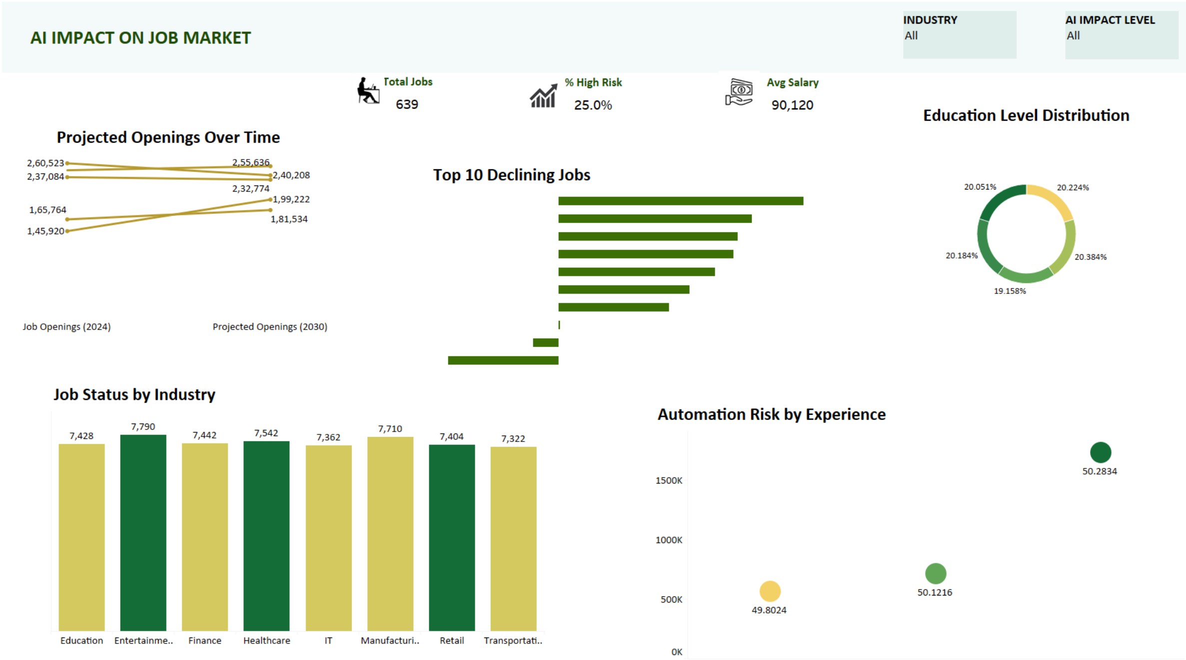
```r
# install package
install.packages("dplyr")
library(dplyr)


# load data set
df <- read.csv("C:\\Users\\USER\\OneDrive\\Documents\\final project\\ai_job_trends_dataset.csv")
View(df)

head(df)
str(df)
summary(df)
```

```r
# T-Test (Compare two group means)
# Test whether Median Salary differs significantly between High vs Low Gender Diversity groups

# Create groups based on Gender Diversity median
med <- median(df$Gender.Diversity...., na.rm = TRUE)
group <- ifelse(df$Gender.Diversity.... > med, "High", "Low")
# Run t-test
result <- t.test(df$Median.Salary..USD. ~ group)
print(result)
# Print simplified interpretation
if(result$p.value < 0.05) {
  print("The difference is statistically significant( p < 0.05\n")
} else {
  print("The difference is not statistically significant (p > 0.05\n")
}
```

```r
# F-Test (Compare variances of two groups)
#  Do high-risk and low-risk job groups have different variance in salary?
# Create a RiskGroup column based on Automation Risk threshold
df$RiskGroup <- ifelse(df$`Automation.Risk....` > 70, "High Risk", "Low/Moderate Risk")
# Now run the F-test
result2<-var.test(`Median.Salary..USD.` ~ RiskGroup, data = df)
print(result2)
if(result2$p.value < 0.05) {
  print("The difference is statistically significant( p < 0.05\n")
} else {
  print("The difference is not statistically significant (p > 0.05\n")
}


# Z-Test (Compare mean of one group to a known value)
# Is the average Automation Risk (%) significantly different from 50?
install.packages("BSDA")
library(BSDA)
result3<-z.test(df$`Automation.Risk....`, mu = 50, sigma.x = sd(df$`Automation.Risk....`, na.rm=TRUE))
print(result3)
if(result3$p.value < 0.05) {
  print("The difference is statistically significant( p < 0.05\n")
} else {
  print("The difference is not statistically significant (p > 0.05\n")
}
```

```r
# ANOVA (Compare means across 3 or more groups)
# checks if the mean number of job openings (2024) differs significantly across different levels of required education

anova_result <- aov(`Job.Openings..2024.` ~ `Required.Education`, data = df)
summary_result <- summary(anova_result)
pval <- summary_result[[1]][["Pr(>F)"]][1]

if (pval < 0.05) {
  print("significant: difference in job openings across different education levels in 2024")
} else {
  print("Not significant")
}
# There is a statistically significant difference in job openings across different education levels in 2024.



# Chi-Square Test
# checks whether there is a significant association between the type of industry and the education level required.
chisq_result1 <- chisq.test(table(df$Industry, df$Required.Education))
print(chisq_result1)
if(chisq_result1$p.value < 0.05) {
  print("The difference is statistically significant( p < 0.05\n")
} else {
  print("The difference is not statistically significant (p > 0.05\n")
}
#  There is a significant relationship between Industry and Required Education
```

Tableau


AI IMPACT ON JOB MARKET
INDUSTRY
All
AI IMPACT LEVEL
All
Total Jobs
639
% High Risk
25.0%
Avg Salary
90,120
Projected Openings Over Time
2,60,523
2,37,084
1,65,764
1,45,920
2,55,636
2,40,208
2,32,774
1,99,222
1,81,534
Job Openings (2024)
Projected Openings (2030)
Top 10 Declining Jobs
Education Level Distribution
20.051%
20.224%
20.184%
20.384%
19.158%
Job Status by Industry
7,428
7,790
7,442
7,542
7,362
7,710
7,404
7,322
Education
Entertainme..
Finance
Healthcare
IT
Manufacturi..
Retail
Transportati..
Automation Risk by Experience
1500K
1000K
500K
0K
49.8024
50.1216
50.2834