Unemployment Analysis ( The Final Project)

**Purpose of the Project**

* To identify key factors influencing unemployment across demographics, industries, and education levels.
* To analyse how variables like experience, education, and AI exposure affect employability and salary.
* To support data-driven workforce planning and policy recommendations.
* To uncover patterns in job application behaviour and platform usage.
* To evaluate the role of AI risk perception in employment outcomes.
* To bridge the gap between raw data and strategic insights.
* To contribute to building a more inclusive and responsive employment ecosystem.

**Nature of the Dataset**

* Contains individual-level data including age, education, job title, industry, salary, employment status, and application platform.
* Includes both categorical (e.g., education level, industry) and numerical (e.g., salary, age) variables.
* Structured and clean, suitable for regression, ANOVA, and Chi-square testing.
* Reflects real-world diversity in employment conditions and applicant behavior.

**Type of Project**

A **descriptive and inferential analytics project** focused on labor market dynamics.

* Combines exploratory data analysis with statistical testing to uncover relationships.
* Hybrid in nature — blending business intelligence with academic rigor.
* Designed to support strategic decision-making and portfolio demonstration.
* Includes visualizations, hypothesis testing, and regression modeling.
* Output is tailored for dashboards, executive summaries, and stakeholder presentations.

**Tests Involved:**

* 1. **Regression Analysis: -**
* Used to quantify relationships between independent variables (e.g., education, experience) and outcomes like salary or employability.
* Helps determine whether these factors significantly predict employment status or income level.
* Reveals which variables have measurable impact and which do not.
* Supports predictive modelling and strategic workforce planning.
* Highlights gaps where traditional assumptions may not hold true.
  1. **ANOVA (Analysis of Variance) :-**
* Applied to test whether group differences (e.g., across industries or education levels) significantly affect outcomes.
* Compares means across multiple groups to determine statistical significance.
* Identifies structural disparities in salary, employability, or AI risk perception.
* Validates whether categorical factors influence employment metrics.
* Guides targeted interventions like reskilling programs or platform-specific outreach.

Let me know if you'd like this formatted into a dashboard narrative, HTML portfolio section, or stakeholder-ready report.

Data Interpretation :-

1. The regression and ANOVA analyses reveal that neither years of experience nor education level significantly influence monthly salary in this dataset, suggesting other factors may be more critical in salary determination.
2. Education strongly impacts employability, indicating that investing in educational programs and upskilling can substantially improve job prospects.
3. Age shows a significant association with education level, which may reflect career progression stages or access to educational opportunities over time.
4. The Chi-square test for AI risk perception across industries helps identify whether risk concerns vary by sector, guiding targeted risk management and reskilling efforts.
5. Including or excluding certain data columns does not affect the dependent variable in the tested models, implying data selection should focus on relevance rather than quantity.

**REGRESSION & ANOVA INTERPRETATIONS**

**1. Monthly Salary vs Years of Experience**

* **F-statistic**: 1.797
* **Significance (p-value)**: 0.1868
* **Conclusion**: Not statistically significant (p > 0.05)
* **Interpretation**: Experience does **not** significantly impact salary in this dataset.

**2. Monthly Salary vs Education**

* **F-statistic**: 0.0999
* **Significance (p-value)**: 0.7526
* **Conclusion**: Not statistically significant
* **Interpretation**: Education level does **not** significantly affect salary.

**3. Employability vs Education**

* **F-statistic**: 83.77
* **Significance (p-value)**: 0.0000
* **Conclusion**: **Highly significant**
* **Interpretation**: Education has a **strong positive impact** on employability.

**4. Education vs Age**

* **F-statistic**: 21.88
* **Significance (p-value)**: 0.000004
* **Conclusion**: **Highly significant**
* **Interpretation**: Age is **strongly associated** with education level — possibly reflecting career stage or access to education.

**CHI-SQUARE TEST INTERPRETATION**

**AI Risk vs Industry**

* **P-value**: Not explicitly shown, but based on expected structure and totals, you would compute it using:
  + Observed vs Expected frequencies
  + Degrees of freedom: (rows - 1) × (columns - 1) = (7 - 1) × (3 - 1) = 12
* If p-value < 0.05 → AI risk perception **varies significantly** by industry
* If p-value ≥ 0.05 → AI risk perception is **uniform across industries**

**ANOVA: Columns Included vs Excluded**

**Both Tests:**

* **F-statistic**: 3.17207E-17
* **P-value**: 1
* **Conclusion**: No significant difference between groups
* **Interpretation**: Including or excluding columns has **no measurable impact** on the dependent variable in this test.

**PROFESSIONAL RECOMMENDATIONS**

**Education Strategy**

* **Invest in education** to boost employability — strong statistical support.
* Consider targeted upskilling programs for younger age groups to align with educational trends.

**Salary Modeling**

* **Avoid over-reliance on education or experience** as salary predictors — neither showed significance.
* Explore other variables like **industry, job role, or skill proficiency** for better salary modeling.

**AI Risk Management**

* If Chi-square test shows significance:
  + Tailor **AI risk mitigation strategies** by industry.
  + Prioritize **reskilling** in high-risk sectors.
* If not significant:
  + Treat AI risk perception as **uniform**, and apply broad awareness campaigns.