**Name Mamoon Haseeb Khan**

**Muhammad Mursaleen**

**Phase 1: Project Proposal**

**Dataset Title:** Upwork Jobs 2024

**Dataset Source:** Kaggle  
The dataset was downloaded from Kaggle and is in CSV (Comma-Separated Values) format.

**Description:**  
This dataset contains job listings posted on Upwork in the year 2024. It includes details such as job title, category, required skills, budget, and number of proposals. It gives a good overview of the freelance market on Upwork.

**Research Question:**  
Can we build a model that predicts how popular or in-demand a job is?  
We want to find out which skills are most demanded on Upwork, so that new freelancers can choose the right skills to learn based on job trends and popularity.

**Potential Challenges and Solutions:**

* **Missing or unclear data:** Some job listings might not have complete information.  
  *Solution:* We will clean the dataset and remove or fill in missing values where possible.
* **Too many different skill names:** Similar skills might be written in different ways.  
  *Solution:* We will group similar skills and use data preprocessing techniques to simplify them.
* **Unbalanced data:** Some job categories might have more data than others.  
  *Solution:* We will use sampling or other techniques to balance the dataset during training.

**Goal:**  
The final goal is to build a machine learning model that can help new freelancers see which skills have the most job opportunities, so they can make better career choices.