### **Section 2.1 Homework: Data Wrangling and Cleaning with Glassdoor Dataset**

#### **Objective: Use the** [**Glassdoor dataset**](https://www.kaggle.com/datasets/rashikrahmanpritom/data-science-job-posting-on-glassdoor?select=Uncleaned_DS_jobs.csv) **to practice data cleaning and wrangling techniques in Excel.**

### **Part 1: Data Cleaning**

#### **Handle Missing Data:**

* Check for missing values in columns like Salary Estimate, Location, and Job Title using conditional formatting (**Home** → **Conditional Formatting** → **Highlight Cell Rules** → **Blanks**).
* Replace missing values:
  + For numerical columns like Salary Estimate: Replace with the average using =AVERAGE(range).
  + For categorical columns like Location: Replace with the most frequent value using =MODE(range).

#### **Remove Duplicates:**

* Use Excel’s **Remove Duplicates** feature (**Data** → **Remove Duplicates**) to identify and remove duplicate rows.

#### **Standardize Text:**

* For columns like Company Name and Job Title, standardize text:
  + Use =PROPER(range) to capitalize company names properly.
  + Use =LOWER(range) to convert text to lowercase.

### **Part 2: Data Wrangling**

#### **Organize Columns:**

* Ensure columns are clearly labeled and irrelevant ones are removed (e.g., any unnecessary metadata or empty fields).

#### **Verify Data Quality:**

* Check for unique values in the Job Title or Company Name columns using:
  + =COUNTIF(range, value).

### **Part 3: Data Visualization**

#### **Bar Chart:**

* Create a bar chart showing the frequency of the top 5 job titles.

### **Part 4: Advanced Analysis**

#### **Conditional Formatting:**

* Highlight rows where the Salary Estimate exceeds $100,000 using Conditional Formatting.

#### **Top 10 Analysis:**

* Use Excel’s **SORT** function to identify the top 10 companies offering the highest salary ranges.

### **Reflection Questions:**

* What percentage of job postings are missing salary estimates?
* Which job title appears most frequently?
* Which location has the highest number of postings?