CS Rubric – Case Study Replication

DS 4002 Course

This document is a rubric for the replication of the process and results of the case study "*Indeed.com* Job Posting Patterns and Predictions Over Time," project conducted in the DS 4002 course in Spring 2025. Contextual information about the original project and case study can be found in the Hook document in this folder, which also contains the linked GitHub repository.

How will you reach success? You will meet spec on CS Case Study Replication by following the steps and criteria below to produce the necessary outputs.

Spec Category	Spec Details
Purpose and Task	● Goal: To replicate the process and outputs of the case study "*Indeed.com* Job Posting Patterns and Predictions Over Time," a project completed in the DS 4002 course in Spring 2025 by running the scripts in the repository SCRIPTS folder and producing the graphical outputs, as well as a document containing a description of the outputs and relevant conclusions that can be drawn. ● Main components: ○ Replicated Exploratory Data Analysis (EDA) plots (.png images) ○ Replicated analysis output plots (.png images) ○ Description of outputs and conclusions drawn (PDF document)
Replicated EDA Plots	● Goal: Replicate primary EDA plots to draw some initial conclusions about job sectors with the most postings in the past six months, job postings across all sectors over the past five years, and job postings organized by U.S. states. ● Install necessary packages stated in the README in your Python environment. ● Format: .png images ● Plots to replicate: ○ sector_summary.png ▪ Follow along and run "dataAnalysis.py" from SCRIPTS folder utilizing "job_postings_by_sector_US (3).csv" from the DATA folder. ○ sector_line_6month.png

	▪ Follow along and run “dataAnalysis01.py” from SCRIPTS folder utilizing “job_postings_by_sector_US (3).csv” from the DATA folder. ○ us_map_states.png ▪ Follow along and run “dataAnalysis01.py” from SCRIPTS folder utilizing “state_job_postings_us (1).csv” from the DATA folder.
Replicated Analysis Output Plots	● Goal: Replicate SARIMAX forecasting plots to draw some final conclusions about the job postings across all sectors in the next five years, job postings organized by the top ten sectors over the next five years, and job postings organized by the top ten U.S. metro cities over the next five years. ● Install necessary packages stated in the README in your Python environment. ● Format: .png images ● Plots to replicate: ○ job_postings_time_series_forecast.png ▪ Follow along and run “aggregateForecast.py” from the SCRIPTS folder utilizing “aggregate_job_postings_US (2).csv” from the DATA folder. ○ metro_forecasts.png ▪ Follow along and run “metroForecast.py” from the SCRIPTS folder utilizing “metro_job_postings_us (1).csv” from the DATA folder. ○ top_sector_forecasts.png ▪ Follow along and run “sectorForecast.py” from the SCRIPTS folder utilizing “job_postings_by_sector_US (3).csv” from the DATA folder.
Description of Outputs and Conclusions Drawn	● Goal: Create a document with the .png images from the EDA and SARIMAX output plots as well as descriptions of their results and relevant conclusions that can be drawn in relation to the U.S. job market on *Indeed.com* that will be helpful for potential job seekers. ● Format: PDF document ● Components ○ .png images of EDA and SARIMAX analysis plots created ○ Written descriptions of the results of each

	○ Relevant conclusions for potential job seekers

Acknowledgements: Special thanks to Jess Taggart from UVA CTE for coaching on making this rubric. This structure is pulled from Streifer & Palmer (2020).