DOUGLAS COLLEGE

COMMERCE & BUSINESS ADMINISTRATION
COMPUTING STUDIES & INFORMATION SYSTEMS
COMPUTER AND INFORMATION SYSTEMS (PBD)

CSIS 4495-050: APPLIED RESEARCH PROJECT

Progress Report 5:
End-to-End Data Engineering Solution for HR Analytics

Student Name: Jay Clark Bermudez | 300380540

Instructor: Dr. Bambang Sarif

NEW WESTMINSTER/BC
FALL/2025

1.0 Work Hours

Date	Number of Hours	Description of Work Done
11/08/2025	5	Finish recording the midterm video
11/09/2025	1.5	Resolve git issues regarding secret code on web app • Added env loading, environment variable checks, and a reusable load_hr_data() function
11/09/2025	2.5	Optimize app design and changed CSS styling • Defined a single global df loaded once at startup and reused in all callbacks.
11/15/2025	3	Convert notebook code to be compatible with venv • Exported the Dash app from ipynb into a standalone python script so it can run using the command line instead of Jupyter
11/16/2025	3	Test and fix bugs on venv • Remove or redo notebook exclusive codes • Trendline in graph 3 is not working
11/18/2025	2	Update inconsistencies for web app documentation due to changes after midterm
11/20/2025	1	Create progress report 5

2.0 Description of Work Done

Between November 8 and November 20, 2025, our team made several updates to the HR analytics dashboard.

On November 8, 2025, we finished recording the midterm video to document the current features and data pipeline. On November 9, 2025, I resolved the git issue related to exposed secrets in the web app by moving credentials into a .env file, adding environment variable checks, and creating a reusable load_hr_data() function to pull cleaned data from Databricks; on the same day, I also optimized the app's visual design by adjusting the CSS styling and refactoring the code so a single global dataframe is loaded once at startup and reused across all callbacks.

On November 15, 2025, I converted the notebook-based prototype into a standalone Python script that runs in a virtual environment using python app.py, and on November 16, 2025, I tested the new setup by removing or rewriting notebook-exclusive code, such as Jupyter-specific flags and run settings, so the app could run cleanly on its own.

During that testing on November 16, I fixed the trendline issue in the 3rd graph (Performance and Experience Relationship), where the OLS regression was not working because the required library was only installed in the original Jupyter environment, and resolved it by installing the necessary package in the venv so the trendline would render correctly.

On November 18, 2025, I updated the web app documentation to reflect all post-midterm changes, and on November 20, 2025, I prepared Progress Report 5 to summarize the recent development work.