



DOUGLAS COLLEGE

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

CSIS 4495-050: APPLIED RESEARCH PROJECT

Progress Report 3:
End-to-End Data Engineering Solution for HR Analytics

Student Name: Bruno do Nascimento Beserra | 300392300

Instructor: Dr. Bambang Sarif

NEW WESTMINSTER/BC
FALL/2025

1.0 Work Hours

Student Name: Bruno do Nascimento Beserra

Date	Number of Hours	Description of Work Done
10/09/2025	0.5	Uploaded code of initial dataset creation
10/11/2025	2	Started working on snapshots logic
10/12/2025	4	Finished snapshots logic and created 84 snapshots of data
10/15/2025	2.5	Created coding for bronze layer of the data engineering pipeline
10/17/2025	2	Started coding for silver layer of the data engineering pipeline
10/18/2025	2.5	Finished working in the silver layer of the data engineering pipeline
10/19/2025	0.5	Updated description of bronze and Silver notebooks
10/19/2025	1.5	Helped Matheus developing the gold layer of the data engineering pipeline
10/22/2025	3.5	Add and update Dataset Creation and Data Engineering parts for Bronze and Silver in Midterm Report
10/23/2025	1	Create Progress Report 2 Document

2.0 Description of Work Done

Between October 9th and October 23rd, 2025, my work focused on advancing the implementation of our data engineering pipeline and enhancing project documentation. I began by creating and uploading the initial dataset, establishing the foundation for future processing and transformation tasks. Following that, I developed the snapshot logic to create 84 snapshots representing different states of the data over time.

Once the snapshots were completed, I moved on to building the bronze layer of the data pipeline, where I implemented the ingestion and initial transformation processes to convert raw data into a structured, accessible format. Afterward, I focused on the silver layer, which involved cleaning, joining, and refining the data to ensure consistency, accuracy, and usability for analytical purposes.

Then, I updated the documentation for both the bronze and silver notebooks. And, I worked with Matheus on the development of the gold layer to make it able to fit the pipeline already created, making the data ready for business insights and reporting.

Finally, I update the Midterm Report, adding detailed descriptions of the dataset creation, bronze, and silver layers, completed the Progress Report 2 document, summarizing all tasks and progress achieved to date, and exported all project files to GitHub to check-in the work done in the repository. Both the dataset and the corresponding Python notebooks were organized and uploaded under the project's Implementation folder.