

Group 4						
Vanessa Gleason, Alistair Marsden, Olivier Rochaix, Eduard Stalmakov						
Capstone Project						
Project Management Plan						
Milestone	Task	Assigned to	Estimated time (total hrs)	Estimated Start	Estimated Finish	Task Status
Capstone Outline						
	Create Github repository for collaboration	Eduard	0.5	9/16/22	9/16/22	Complete
	Start and structure a README file for GitHub	Eduard	1.5	9/16/22	9/16/22	Complete
	Decide on exploratory questions	All	24	9/20/22	9/20/22	Complete
	Stored on GitHub as "ExploratoryQuestions.pdf"	Eduard	0.5	9/20/22	9/20/22	Complete
	Find datasets [user playlists, song attributes, and user demographics]	All	48	9/20/22	9/20/22	Complete
	Research demographic data	Vanessa	6	9/20/22	9/20/22	Complete
	Find census datasets [Age, Income, Computer Access]	Alistair/Vanessa	6	9/23/22	9/23/22	Complete
	Research machine learning algorithm	All	48	9/21/22	9/21/22	Complete
	Create elevator pitch formal summary	All	8	9/20/22	9/20/22	Complete
	Complete initial PMP	All	16	9/22/22	9/22/22	Complete
	Stored on GitHub as "ProjectManagementPlan.pdf"	Vanessa	0.5	9/22/22	9/22/22	Complete
	Create ER diagram	Eduard	2	9/21/22	9/21/22	Complete
	Data platform napkin drawing	Olivier	1	9/20/22	9/20/22	Complete
	Create ETL document structure	Vanessa	0.5	9/22/22	9/22/22	Complete
	Stored on GitHub as "RepeatableETLReport.pdf"	Vanessa	0.5	9/22/22	9/22/22	Complete
	Create list of all data sources used for the project	Alistair	1	9/23/22	9/23/22	Complete
	Stored on GitHub as "DataSources.pdf"	Alistair	1	9/23/22	9/23/22	Complete
Submit on Github						
Access						
	Create SQL database	Eduard	0.5	9/20/22	9/20/22	Complete
	Create group container	Eduard	0.5	9/20/22	9/20/22	Complete
	Create Kafka topic	Eduard	0.5	9/20/22	9/20/22	Complete
	Create a Data Factory	Eduard	1	9/20/22	9/20/22	Complete
	Create Python environment	Vanessa	2	9/26/22	9/26/22	Complete
	Update Python environment for Dash	Vanessa	1	9/27/22	9/27/22	Complete
Produce and Consume Kafka Data using Data Factory						
	Create producer in DataBricks	Eduard	1	9/22/22	9/22/22	Complete

Group 4						
Vanessa Gleason, Alistair Marsden, Olivier Rochaix, Eduard Stalmakov						
Capstone Project						
Project Management Plan						
Milestone	Task	Assigned to	Estimated time (total hrs)	Estimated Start	Estimated Finish	Task Status
	Read last.fm songs into topic with producer	Eduard	1	9/22/22	9/22/22	Complete
	Set up data factory pipeline from data set to the producer in Data Factory	Eduard	1	9/22/22	9/22/22	Complete
	Create consumer in DataBricks to collect messages from the Kafka server	Eduard	1	9/22/22	9/22/22	Complete
Reading Kafka Files from Data Lake Into the Database						
	Create one or more tables for the finished messages you are consuming	Alistair	1	9/22/22	9/22/22	Complete
	Set up data factory pipeline from from the producer to the consumer	Eduard	1	9/22/22	9/22/22	Complete
	Export the data to the group database	Eduard	1	9/22/22	9/22/22	Complete
Diagram and Document Your Cloud Servers						
	Detail and document the set of tools used / planned for use (Initial Draft)	Olivier	3	9/22/22	9/22/22	Complete
	Stored on GitHub as "DataPlatformDiagram.pdf"	Olivier	0.5	9/22/22	9/22/22	Complete
	Performed revisions on Data Platform Diagram and documentation	Olivier	3	9/25/22	9/25/22	Complete
	Additional revisions	Olivier	1	9/28/22	9/28/22	Complete
Submit on Github						
ETL						
	Collect all raw data files in appropriate container folder	Eduard	1	9/23/22	9/23/22	Complete
	Perform ETL on last.fm data	Eduard	3	9/23/22	9/26/22	Complete
	Perform EDA on last.fm data	Eduard	6	9/23/22	9/26/22	Complete
	Perform ETL on songs data	Olivier	6	9/23/22	9/26/22	Complete
	Perform EDA on songs data	Olivier	3	9/23/22	9/26/22	Complete
	Perform ETL on census data	Alistair	3	9/23/22	9/26/22	Complete
	Perform EDA on census data	Alistair	6	9/23/22	9/26/22	Complete
	Store refined data files in appropriate container folder	All	4	9/23/22	9/26/22	Complete
Normalization and SQL Database						
	Reach consensus on and finalize ERD	All	12	9/26/22	9/27/22	Complete
	Create tables matching ERD	All	12	9/26/22	9/27/22	Complete

Group 4						
Vanessa Gleason, Alistair Marsden, Olivier Rochaix, Eduard Stalmakov						
Capstone Project						
Project Management Plan						
Milestone	Task	Assigned to	Estimated time (total hrs)	Estimated Start	Estimated Finish	Task Status
	Insert last.fm data into table	Eduard	3	9/26/22	9/27/22	Complete
	Insert songs data into table	Olivier	6	9/26/22	9/27/22	Complete
	Insert census data into table	Alistair	5	9/26/22	9/27/22	Complete
	Create pipelines with ETL DataBricks	Eduard	2	9/27/22	9/27/22	Complete
Machine Learning Model						
	Research/trial collaborative vs content ML	Vanessa/Eduard	6	9/26/22	9/26/22	Complete
	Create an initial collaborative recommender model	Vanessa	5	9/26/22	9/27/22	Complete
	Create an initial content based recommender model	Eduard	5	9/27/22	9/27/22	Complete
	Integrate SQL data into ML model	Vanessa/Eduard	10	9/29/22	9/29/22	Complete
	Complete collaborative recommender model	Vanessa	8	9/29/22	9/29/22	Complete
	Stored on GitHub as "CollaborativeRecommender.py"	Vanessa	0.5	9/29/22	9/29/22	Complete
	Complete content based recommender model	Eduard	6	9/29/22	9/29/22	Complete
	Stored on GitHub as "Content-Based-Recommender.py"	Eduard	0.5	9/29/22	9/29/22	Complete
Submit on Github						
Napkin Drawings and Feedback for Visualizations						
	Create Napkin Drawing(s)	All	24	9/28/22	9/28/22	Complete
	Formalize napkin drawings and share for feedback	Olivier	1	9/29/22	9/29/22	Complete
	Document Feedback	Alistair	2	9/29/22	9/29/22	Complete
	Stored on GitHub as "DashboardNapkinsAndFeedback.pdf"	Olivier	0.5	9/29/22	9/29/22	Complete
Dashboard						
	Create Dashboard skeleton	Vanessa	3	9/27/22	9/27/22	Complete
	Stored on GitHub in "Dashboard" folder	Vanessa	1	9/29/22	9/29/22	Complete
	Add data sources to Dashboard	Eduard	1	9/29/22	10/2/22	Complete
	Choose color theme	All	4	9/29/22	10/3/22	Complete
	Make visuals (barchart, graph, heat map, pie, etc)	Vanessa	3	9/29/22	10/2/22	Complete
	Make visuals (barchart, graph, heat map, pie, etc)	Alistair	3	9/29/22	10/2/22	Complete
	Make visuals (barchart, graph, heat map, pie, etc)	Eduard	3	9/29/22	10/2/22	Complete
	Make visuals (barchart, graph, heat map, pie, etc)	Olivier	3	9/29/22	10/3/22	Complete

Group 4						
Vanessa Gleason, Alistair Marsden, Olivier Rochaix, Eduard Stalmakov						
Capstone Project						
Project Management Plan						
Milestone	Task	Assigned to	Estimated time (total hrs)	Estimated Start	Estimated Finish	Task Status
	Review and complete Dashboard	All	48	9/30/22	10/5/22	Complete
Submit on Github						
ETL Documentation						
	Fill out group ETL process in ETL document	All	12	10/3/22	10/5/22	Complete
	Review ETL document and repeat	All	4	10/3/22	10/5/22	Complete
	Stored on GitHub as "RepeatableETLReport.pdf"	Olivier	1	10/5/22	10/5/22	Complete
Project Technical Report						
	Create report with initial structure	Vanessa	1	9/28/22	9/28/22	Complete
	Additional research	All	8	10/3/22	10/5/22	Complete
	Fill in body of report	All	12	10/3/22	10/5/22	Complete
	Review Report and Complete	All	8	10/3/22	10/5/22	Complete
	Stored on Github as "ProjectTechnicalReport.pdf"	Vanessa	0.5	10/3/22	10/5/22	Complete
Presentation Slides						
	Create Google Slides with initial structure	Alistair	0.5	10/4/22	10/5/22	Complete
	Enter information into slides	All	12	10/4/22	10/6/22	Complete
	Review and complete presentation	All	8	10/4/22	10/6/22	Complete
	Stored on GitHub as "PresentationSlides.pdf"	Alistair	0.5	10/4/22	10/6/22	Complete
Presentation Dry Run						
	Practice Presenting	All	2	10/6/22	10/6/22	Complete
Submit on Github						