

Cohort 31 - Group 4							Name	Number of	Task Hours
Vanessa Gleason, Alistair Marsden, Olivier Rochaix, Eduard Stalmakov							Eduard	21	31
Capstone Project							Vanessa	9	19.5
Project Management Plan							Alistair	10	14.5
Milestone	Task	Assigned to	Estimated time (hrs)	Estimated Start	Estimated Finish	Task Status	Olivier	10	27.5
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	6	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	12	9/20/22	9/20/22	Complete			
	Find a dataset to link census data to audio streaming market	All	6	9/23/22	9/23/22	IP			
	Find census datasets [Age, Income, Computer Access]	All	6	9/23/22	9/23/22	IP			
	Research machine learning algorithm	All	12	9/21/22	9/21/22	Complete			
	Create elevator pitch formal summary	All	2	9/20/22	9/20/22	Complete			
	Complete initial PMP	All	4	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	4	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	IP			
	Stored on GitHub as "DataSources.pdf"	Alistair	1	9/23/22	9/23/22	IP			
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			
Produce and Consume Kafka Data using Data Factory									
	Create producer in DataBricks	Eduard	1	9/22/22	9/22/22	Complete			
	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									

Cohort 31 - Group 4							Name	Number of	Task Hours
Vanessa Gleason, Alistair Marsden, Olivier Rochaix, Eduard Stalmakov							Eduard	21	31
Capstone Project							Vanessa	9	19.5
Project Management Plan							Alistair	10	14.5
Milestone	Task	Assigned to	Estimated time (hrs)	Estimated Start	Estimated Finish	Task Status	Olivier	10	27.5
	Detail and document the set of tools used / planned for use	Olivier	1	9/22/22	9/22/22	Complete			
	Stored on GitHub as "DataPlatform.pdf"	Olivier	0.5	9/22/22	9/22/22	Complete			
Submit on Github									
ETL									
	Collect all raw data files in appropriate container folder	Eduard	1	9/23/22	9/23/22	IP			
	Perform ETL on last.fm data	Eduard	3	9/23/22	9/26/22	IP			
	Perform EDA on last.fm data	Eduard	6	9/23/22	9/26/22	IP			
	Perform ETL on songs data	Olivier	3	9/23/22	9/26/22	IP			
	Perform EDA on songs data	Olivier	6	9/23/22	9/26/22	IP			
	Perform ETL on census data	Vanessa / Alistair	3	9/23/22	9/26/22	IP			
	Perform EDA on census data	Vanessa / Alistair	6	9/23/22	9/26/22	IP			
	Store refined data files in appropriate container folder	All	1	9/23/22	9/26/22	IP			
Normalization and SQL Database									
	Reach consensus on and finalize ERD	All	3	9/26/22	9/26/22	NS			
	Create tables matching ERD	Alistair	3	9/26/22	9/26/22	NS			
	Insert last.fm data into table	Eduard	2	9/26/22	9/26/22	NS			
	Insert songs data into table	Olivier	2	9/26/22	9/26/22	NS			
	Insert census data into table	Vanessa	2	9/26/22	9/26/22	NS			
Machine Learning Model									
	Prepare Tracks table for chosen ML model	Eduard	2	9/27/22	9/27/22	NS			
	Create a KNN model using data from the prepared Tracks table	Vanessa	6	9/27/22	9/27/22	NS			
	Prepare Playlists table for chosen ML model	Alistair	2	9/27/22	9/27/22	NS			
	Create a second ML model using data from the prepared Playlists table	Olivier	6	9/27/22	9/27/22	NS			
	Compare Playlist model results to Tracks model results to get recommendations	Vanessa	6	9/27/22	9/27/22	NS			
Submit on Github									
Napkin Drawings and Feedback for Visualizations									
	Create Napkin Drawing(s)	All	6	9/28/22	9/28/22	NS			
	Stored on GitHub as "VisualizationsNapkinsAndFeedback.pdf"	All	0.5	9/28/22	9/28/22	NS			
	Document Feedback	Alistair	2	9/28/22	9/28/22	NS			
	Stored on GitHub as "DashboardNapkinsAndFeedback.pdf"	Alistair	0.5	9/28/22	9/28/22	NS			
Dashboard									
	Create Dashboard	Olivier	1	9/29/22	9/29/22	NS			
	Stored on GitHub as "Dashboard.pdf"	Olivier	1	9/29/22	9/29/22	NS			
	Add data sources to Dashboard	Eduard	1	9/29/22	9/29/22	NS			
	Choose color theme	All	1	9/29/22	9/29/22	NS			
	Make visuals (barchart, graph, heat map, pie, etc)	Vanessa	3	9/29/22	9/29/22	NS			

Cohort 31 - Group 4							Name	Number of	Task Hours
Vanessa Gleason, Alistair Marsden, Olivier Rochaix, Eduard Stalmakov							Eduard	21	31
Capstone Project							Vanessa	9	19.5
Project Management Plan							Alistair	10	14.5
Milestone	Task	Assigned to	Estimated time (hrs)	Estimated Start	Estimated Finish	Task Status	Olivier	10	27.5
	Make visuals (barchart, graph, heat map, pie, etc)	Alistair	3	9/29/22	9/29/22	NS			
	Make visuals (barchart, graph, heat map, pie, etc)	Eduard	3	9/29/22	9/29/22	NS			
	assign/make visuals (barchart, graph, heat map, pie, etc)	Olivier	3	9/29/22	9/29/22	NS			
	Complete Dashboard	All	6	9/30/22	9/30/22	NS			
Submit on Github									
Project Technical Report									
	Create report with initial structure	Vanessa	0.5	10/3/22	10/3/22	NS			
	Additional research	All	2	10/3/22	10/3/22	NS			
	Fill in body of report	All	3	10/3/22	10/3/22	NS			
	Review Report and Complete	All	2	10/3/22	10/3/22	NS			
	Stored on Github as "ProjectTechnicalReport.pdf"	Vanessa	0.5	10/3/22	10/3/22	NS			
Presentation Slides									
	Create Google Slides with initial structure	Alistair	0.5	10/4/22	10/4/22	NS			
	Enter information into slides	All	3	10/4/22	10/4/22	NS			
	Review and complete presentation	All	2	10/4/22	10/4/22	NS			
	Stored on GitHub as "PresentationSlides.pdf"	Alistair	0.5	10/4/22	10/4/22	NS			
Presentation Dry Run									
	Practice Presenting	All	0.5	10/4/22	10/4/22	NS			
Submit on Github									
Stretch Goals									