



Finance Capstone Dev10 ★ ...

[List](#)[Board](#)[Table](#)[Dashboard](#)[Files](#)[Time](#)[Milestones](#)[Forms](#)[More...](#)

Capstone

[+ Add a Task](#)[Automate](#)

- ☒

 Create Repository [More](#) Completed Due Fri Sep 16th Est: 10m
- ☒

 Brainstorm Questions; Search for Datasets; EDA Completed Due Yesterday Est: 5h
- ☒

 Executive Summary - First Draft Completed Today - Today Est: 1h
- ☒

 Finish PMP In progress Today - Today Est: 2h
- ☒

 Finalize Datasets In progress Today - Today Est: 2h
- ☒

 EDA In progress Today - Today Est: 2h
- ☒

 Written Executive Summary (1 paragraph) [More](#) Completed Today - Today Est: 2h
- ☒














 Exploratory Questions pdf (5-10 finalized questions) In progress Due Today Est: 1h 60%
- ☒

 Create two group Containers & SQL Database [More](#) In progress Due Today Est: 1h 50%
- ☒


 Build Producer and Topic [More](#) In progress Tomorrow - Tomorrow Est: 2h
- ☒

 Upload all data sets to azure container In progress Tomorrow - Tomorrow Est: 30m
- ☒

 Create ERD To do Tomorrow - Tomorrow Est: 2h 30m

- ☒  Napkin drawings data pipeline diagram [More](#) ● To do Due Tomorrow
⌚ Est: 1h 30m !
- ☒  Research Best Machine Learning Model for Data [More](#) ● To do Due Tomorrow
- ☒  Create tables in Database based on ERD schema ● To do Due Thu Sep 22nd
- ☒  Build Consumer & Append to SQL database table [More](#) ● To do
Due Fri Sep 23rd ⌚ Est: 2h
- ☒  Final data pipeline diagram [More](#) Due Mon Sep 26th ⌚ Est: 3h !
- ☒  Napkin drawings dashboard + post for feedback [More](#) Due Mon Sep 26th
⌚ Est: 2h !
- ☒  Data Clean & Transform 2-3 Datasets [More](#) Due Mon Sep 26th ⌚ Est: 3h !
- ☒  Create Data Factory & Run Producer/Consumer Automated [More](#)
Due Mon Sep 26th ⌚ Est: 3h !
- ☒  Add rest of cleaned data to SQL database (if not through producer/co... [More](#)
Due Mon Sep 26th ⌚ Est: 3h !
- ☒  ETL Report Introduction & Sources (APA Format) Due Tue Sep 27th
⌚ Est: 1h 30m !
- ☒  ETL Report: Extract, Transform, Load for 2-3 Datasets YOU Cleaned [More](#)
Due Tue Sep 27th ⌚ Est: 2h !
- ☒  Finalized napkin drawing for dashboard [More](#) Due Wed Sep 28th ⌚ Est: 2h
- ☒  Build, Export, & Load ML Model [More](#) Due Wed Sep 28th ⌚ Est: 3h !


- 




Edit ETL Report & Write Conclusion (make it perfect :) [More](#)

Due Wed Sep 28th

⌄ Est: 2h 10m



- 



Create napkin drawings for visualizations (1/2--maybe federal level q... [More](#)

Due Thu Sep 29th

⌄ Est: 2h



- 



Create napkin drawings for visualizations (1/2--state-level questions) [More](#)

Due Thu Sep 29th

⌄ Est: 2h



- 




Refine/hyper-parameter tune MI Model [More](#)

Due Fri Sep 30th

⌄ Est: 3h



- 



Using feedback, create final napkin drawings for ALL visualizations [More](#)

Due Sat Oct 1st

⌄ Est: 2h



- 



Create ReadMe [More](#)

Due Mon Oct 3rd

⌄ Est: 3h



- 



Create Visualizations for State Level Questions using Dash Format [More](#)

Due Mon Oct 3rd

⌄ Est: 3h



- 





Create visualizations for national level questions using Dash format [More](#)

Due Mon Oct 3rd

⌄ Est: 3h



- 




Create visualizations for MI model: predictions, accuracy, etc. (dash f...

Due Mon Oct 3rd

⌄ Est: 2h




- 



Create Dash Dashboard (copy & paste visualizations codes together), ...

Due Tue Oct 4th

⌄ Est: 4h




- 




Presentation Slideshow Intro, Initial Questions, Research Process/ED... [More](#)

Due Tue Oct 4th

⌄ Est: 2h




- 



Presentation Slides Machine Learning Use/Tuning/Results/Accuracy, ... [More](#)

Due Tue Oct 4th

⌄ Est: 2h



- 
























Make sure all code is uploaded to repo (and that code is readable): d...

Due Tue Oct 4th

⌄ Est: 1h



-
- ✓  Presentation Dry Run: Make sure we are in 10 mins range; assign slid...
Due Wed Oct 5th ⌚ Est: 1h 
-
- ✓  Technical Report Executive Summary (copy & paste over), Introductio... [More](#)
Due Wed Oct 5th ⌚ Est: 2h 
-
- ✓  Technical Report Research on Federal Questions (Section I) [More](#)
Due Wed Oct 5th ⌚ Est: 3h 
-
- ✓  Technical Report Research Section II: Explain Machine Learning mode...
Due Wed Oct 5th ⌚ Est: 3h 
-
- ✓  Technical Report Research Section III: Answer questions about state (...
Due Wed Oct 5th ⌚ Est: 3h 
-
- ✓  Technical Report: Conclusion, Edit/Review other sections & introducti...
Due Wed Oct 5th ⌚ Est: 2h 
-
- ✓  Class Presentation [More](#) Due Thu Oct 6th ⌚ Est: 20m 
-
- ✓  All project components COMPLETE & PERFECTED [More](#) Due Thu Oct 6th
⌚ Est: 2h 
-
- ✓  All PDFs, code uploaded to GitHub: presentation slides, final report, d... [More](#)
Due Thu Oct 6th ⌚ Est: 3h 
-
- ✓  Update Repo/ReadMe as needed; make sure everything is uploaded ... [More](#)
Due Thu Oct 6th ⌚ Est: 1h 
-
- ✓  Final Presentation [More](#)  Due Fri Oct 7th ⌚ Est: 20m 
-

Capstone — Task list report

Active Tasks

	Task	Start Date	Due Date	Responsible	Created By	Priority	Progress	Status
	Executive Summary - First Draft <div><div></div>Completed</div>	09/20/2022	09/20/2022	Beth Vander Hoek	Shannon Bayless	Medium	0%	Due today (Not started)
	Finish PMP <div><div></div>In progress</div>	09/20/2022	09/20/2022	Beth Vander Hoek	Shannon Bayless	High	0%	Due today (Not started)
	Finalize Datasets <div><div></div>In progress</div>	09/20/2022	09/20/2022	Anyone	Shannon Bayless	High	0%	Due today (Not started)
	EDA <div><div></div>In progress</div>	09/20/2022	09/20/2022	Anyone	Shannon Bayless	High	0%	Due today (Not started)
	Written Executive Summary (1 paragraph) <div><div></div>Completed</div> Executive summary and list of questions	09/20/2022	09/20/2022	Beth Vander Hoek	Beth Vander Hoek	Medium	0%	Due today (Not started)
	Exploratory Questions pdf (5-10 finalized questions) <div><div></div>In progress</div>		09/20/2022	Beth Vander Hoek	Beth Vander Hoek	Medium	60%	Due today (Started)
	Create two group Containers & SQL Database <div><div></div>In progress</div> Container for producer to pull from, container for consumer to send to, database for eventual data platform		09/20/2022	Shannon Bayless	Beth Vander Hoek	Medium	50%	Due today (Started)
	Build Producer and Topic <div><div></div>In progress</div> Build producer and create topic with unique name relevant to dataset producer is sending (i.e. Drug-Overdose, Crime-Rate, etc.). TELL US WHAT TOPIC NAME IS. Producer will send RAW csv data over to consumer under the designated topic name (no cleaning/editing data beforehand)	09/21/2022	09/21/2022	Shannon Bayless	Shannon Bayless	Medium	0%	Upcoming (within 7 days) (Not started)
	Upload all data sets to azure container <div><div></div>In progress</div>	09/21/2022	09/21/2022	Dev10	Shannon Bayless	High	0%	Upcoming (within 7 days) (Not started)
		09/21/2022	09/21/2022	Anyone	Shannon	Medium	0%	Upcoming

	Create ERD <div><div></div>To do</div>				Bayless			(within 7 days) (Not started)
	Napkin drawings data pipeline diagram <div><div></div>To do</div> How cloud/data factory flow will look like		09/21/2022	Anyone	Beth Vander Hoek	Medium	0%	Upcoming (within 7 days) (Not started)
	Research Best Machine Learning Model for Data <div><div></div>To do</div> Research which ML model to use for data		09/21/2022	Beth Vander Hoek	Beth Vander Hoek		0%	Upcoming (within 7 days) (Not started)
	Create tables in Database based on ERD schema <div><div></div>To do</div>		09/22/2022	Anyone	Beth Vander Hoek		0%	Upcoming (within 7 days) (Not started)
	Build Consumer & Append to SQL database table <div><div></div>To do</div> Create consumer using same topic name as producer. Receives data messages then sends data to other group container (from one group container to the other). Data needs to eventually end up in SQL		09/23/2022	Anyone	Beth Vander Hoek		0%	Upcoming (within 7 days) (Not started)
	Final data pipeline diagram Maybe use icons to look pretty like GitHub repo project diabetes example?		09/26/2022	Anyone	Beth Vander Hoek	Medium	0%	Upcoming (within 7 days) (Not started)
	Napkin drawings dashboard + post for feedback What we want dashboard to look like (ROUGH sketch; nothing too detailed--we lose points for that) Post to teams afterwards to get 2 feedback comments (will use to make final draft)		09/26/2022	Anyone	Beth Vander Hoek	Medium	0%	Upcoming (within 7 days) (Not started)
	Data Clean & Transform 2-3 Datasets Every group member should choose 2-3 datasets to clean IN A DATABRICK. Use comments to make clear what you're doing at each step. Take notes on steps so that ETL report is easy to make		09/26/2022	Dev10	Beth Vander Hoek	High	0%	Upcoming (within 7 days) (Not started)
	Create Data Factory & Run Producer/Consumer Automated Use exercise data factory as baseline.		09/26/2022	Anyone	Beth Vander Hoek	Medium	0%	Upcoming (within 7 days) (Not started)
	Add rest of cleaned data to SQL database (if not through producer/consumer data factory)		09/26/2022	Anyone	Beth Vander Hoek	Medium	0%	Upcoming (within 7 days)

	Make sure no repeats of data; only if not already added. Keep track of steps to add for ETL (merging, etc.)							(Not started)
	ETL Report Introduction & Sources (APA Format)		09/27/2022	Anyone	Beth Vander Hoek	Medium	0%	Upcoming (Not started)
	ETL Report: Extract, Transform, Load for 2-3 Datasets YOU Cleaned For datasets you cleaned earlier (every person has 2-3). Be sure to number steps and be clear as possible		09/27/2022	Dev10	Beth Vander Hoek	Medium	0%	Upcoming (Not started)
	Finalized napkin drawing for dashboard Use feedback from Teams chat to make changes as needed to Dashboard napkin drawing. Keep it ROUGH (we lost points if too detailed). Make note if you decided not to use feedback and why		09/28/2022	Anyone	Beth Vander Hoek		0%	Upcoming (Not started)
	Build, Export, & Load ML Model See exercise/demo in applied kafka folder for help		09/28/2022	Anyone	Beth Vander Hoek	Medium	0%	Upcoming (Not started)
	Edit ETL Report & Write Conclusion (make it perfect :)		09/28/2022	Anyone	Beth Vander Hoek	Medium	0%	Upcoming (Not started)
	Create napkin drawings for visualizations (1/2--maybe federal level questions) Post in teams for feedback afterwards; need 2 pieces of feedback. KEEP ROUGH (detailed = lose points)		09/29/2022	Anyone	Beth Vander Hoek	Medium	0%	Upcoming (Not started)
	Create napkin drawings for visualizations (1/2--state-level questions) Post in teams for 2 comments of feedback; coordinate with other visualizations to post at the same time/combine document		09/29/2022	Anyone	Beth Vander Hoek	Medium	0%	Upcoming (Not started)
	Refine/hyper-parameter tune ML Model Make sure model is at optimal performance: feature engineering, tune hyperparameters, etc.		09/30/2022	Anyone	Beth Vander Hoek	Medium	0%	Upcoming (Not started)
	Using feedback, create final napkin drawings for ALL visualizations Use feedback from 2 comments on Teams; make notes as to why not using feedback if so		10/01/2022	Anyone	Beth Vander Hoek	Medium	0%	Upcoming (Not started)
	Create ReadMe ReadMe includes datasets used, links to files in repo, diagram of cloud/data factory, group member names, exploratory questions, how to navigate repo, etc. (see diabetes capstone github for inspo)		10/03/2022	Anyone	Beth Vander Hoek	Medium	0%	Upcoming (Not started)
	Create Visualizations for State Level Questions using Dash Format Visualizations for state questions (not including ML model, prediction, accuracy). Do in Dash format		10/03/2022	Anyone	Beth Vander Hoek	Medium	0%	Upcoming (Not started)
	Create visualizations for national level questions using Dash format		10/03/2022	Anyone	Beth	Medium	0%	Upcoming

	ONLY for national questions, not Model, state questions				Vander Hoek			(Not started)
	Create visualizations for ML model: predictions, accuracy, etc. (dash format if possible)		10/03/2022	Anyone	Beth Vander Hoek	Medium	0%	Upcoming (Not started)
	Create Dash Dashboard (copy & paste visualizations codes together), add layout features & auto-update		10/04/2022	Anyone	Beth Vander Hoek	High	0%	Upcoming (Not started)
	Presentation Slideshow Intro, Initial Questions, Research Process/EDA slides First half of presentation slideshow		10/04/2022	Anyone	Beth Vander Hoek	High	0%	Upcoming (Not started)
	Presentation Slides Machine Learning Use/Tuning/Results/Accuracy, Gov. Policy Recommendations, Concluding Slide (Sources) Second half of slideshow presentation		10/04/2022	Anyone	Beth Vander Hoek	High	0%	Upcoming (Not started)
	Make sure all code is uploaded to repo (and that code is readable): dash visualizations, databricks, etc.		10/04/2022	Dev10	Beth Vander Hoek	High	0%	Upcoming (Not started)
	Presentation Dry Run: Make sure we are in 10 mins range; assign slides to speakers		10/05/2022	Dev10	Beth Vander Hoek	High	0%	Upcoming (Not started)
	Technical Report Executive Summary (copy & paste over), Introduction, Works Cited Page at end (dataset sources; copy & paste from ETL report) DONE by Wednesday lunchtime		10/05/2022	Anyone	Beth Vander Hoek	High	0%	Upcoming (Not started)
	Technical Report Research on Federal Questions (Section I) First Section of technical report; include visualizations (screen shot from dashboard) for answering federal level questions (two questions)		10/05/2022	Anyone	Beth Vander Hoek	High	0%	Upcoming (Not started)
	Technical Report Research Section II: Explain Machine Learning model, predictions, accuracy, include visualizations; answers questions 3/4 (predict poverty level using factors)		10/05/2022	Anyone	Beth Vander Hoek	High	0%	Upcoming (Not started)
	Technical Report Research Section III: Answer questions about state (excluding machine learning); include visualizations		10/05/2022	Anyone	Beth Vander Hoek	High	0%	Upcoming (Not started)
	Technical Report: Conclusion, Edit/Review other sections & introduction		10/05/2022	Anyone	Beth Vander Hoek	High	0%	Upcoming (Not started)
	Class Presentation Presentation for the class!		10/06/2022	Anyone	Beth Vander Hoek	High	0%	Upcoming (Not started)
	All project components COMPLETE & PERFECTED EOD THURSDAY!!! Project is perfect :)		10/06/2022	Anyone	Beth Vander Hoek	High	0%	Upcoming (Not started)
	All PDFs, code uploaded to GitHub: presentation slides, final report, dashboard/visualizations code, ERD, napkin drawings & feedback (Dashboard &		10/06/2022	Anyone	Beth Vander	High	0%	Upcoming (Not

visualizations), ETL report, ReadMe, Azure Studio Links, data pipeline diagram, exploratory questions THURSDAY MORNING/WEDNESDAY NIGHT ALL DONE!!				Hoek			started)
Update Repo/ReadMe as needed; make sure everything is uploaded properly Thursday morning before presentation start		10/06/2022	Beth Vander Hoek	Beth Vander Hoek	Medium	0%	Upcoming (Not started)
Final Presentation <div><div></div>Completed</div> Final presentation to class, attendees, Dev10, Genesis10, shareholders, clients		10/07/2022	Anyone	Beth Vander Hoek	High	0%	Upcoming (Not started)