# Capstone — Task list report

## Active Tasks

| Task | Start Date | Due Date | Responsible | Created By | Priority | Progress | Status | Estimated |
|---|---|---|---|---|---|---|---|---|
| Napkin drawings dashboard + post for feedback  Blocker  ● To do<br><br>BLOCKER FOR DASHBOARD<br>What we want dashboard to look like (ROUGH sketch; nothing too detailed--we lose points for that)<br><br>Post to teams afterwards to get 2 feedback comments (will use to make final draft) | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 2 hours |
| ETL Report Introduction & Sources (APA Format)  ● In progress | | | Beth Vander Hoek | Beth Vander Hoek | Medium | 0% | (Not started) | 1 hour 30 minutes |
| Finalized napkin drawing for dashboard<br><br>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 | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 2 hours |
| Build, Export, & Load ML Model  Blocker  ● To do<br><br>BLOCKER FOR BUILDING MODEL IN DASH<br>See exercise/demo in applied kafka folder for help | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 3 hours |
| Edit ETL Report & Write Conclusion (make it perfect :) | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 2 hours 10 minutes |
| Create napkin drawings for visualizations (1/2--maybe federal level questions)  Blocker<br><br>BLOCKER FOR FINAL NAPKIN DRAWING<br>Post in teams for feedback afterwards; need 2 pieces of feedback. KEEP ROUGH (detailed = lose points) | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 2 hours |
| Create napkin drawings for visualizations (1/2--state-level questions)  Blocker<br><br>BLOCKER FOR FINAL NAPKIN DRAWINGS<br>Post in teams for 2 comments of feedback; coordinate with other visualizations to post at the same time/combine document | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 2 hours |
| Refine/hyper-parameter tune MI Model  Blocker | | | Anyone | Beth Vander | Medium | 0% | (Not started) | 3 hours |

| Task | | | Assignee | Owner | Priority | % | Status | Time |
|---|---|---|---|---|---|---|---|---|
| BLOCKER FOR FINILIZING THE DASHBOARD<br>Make sure model is at optimal performance: feature engineering, tune hyperparameters, etc. | | | | Hoek | | | | |
| Using feedback, create final napkin drawings for ALL visualizations<br><br>Use feedback from 2 comments on Teams; make notes as to why not using feedback if so | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 2 hours |
| Create ReadMe<br><br>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 githib for inspo) | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 3 hours |
| Create Visualizations for State Level Questions using Dash Format <span>Blocker</span><br><br>BLOCKER FOR BUILDING FINAL DASHBOARD<br>Visualizations for state questions (not including ML model, prediction, accuracy). Do in Dash format | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 3 hours |
| Create visualizations for national level questions using Dash format <span>Blocker</span><br><br>BLOCKER FOR BUILDING FINAL DASHBOARD<br>ONLY for national questions, not Model, state questions | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 3 hours |
| Create visualizations for ML model: predictions, accuracy, etc. (dash format if possible) <span>Blocker</span><br><br>BLOCKER FOR BUILDING FINAL DASHBOARD | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 2 hours |
| Create Dash Dashboard (copy & paste visualizations codes together), add layout features & auto-update | | | Anyone | Beth Vander Hoek | High | 0% | (Not started) | 4 hours |
| Presentation Slideshow Intro, Initial Questions, Research Process/EDA slides <span>Blocker</span><br><br>BLOCKER FOR PRESENTATION DRY RUN<br>First half of presentation slideshow | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 2 hours |
| Presentation Slides Machine Learning Use/Tuning/Results/Accuracy, Gov. Policy Recommendations, Concluding Slide (Sources) <span>Blocker</span><br><br>BLOCKER FOR PRESENTATION DRY RUN<br>Second half of slideshow presentation | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 2 hours |
| Make sure all code is uploaded to repo (and that code is readable): dash visualizations, databricks, etc. | | | Dev10 | Beth Vander Hoek | High | 0% | (Not started) | 1 hour |
| Presentation Dry Run: Make sure we are in 10 mins range; assign slides to speakers | | | Dev10 | Beth Vander Hoek | High | 0% | (Not started) | 1 hour |
| Technical Report Executive Summary (copy & paste over), | | | Anyone | Beth | Medium | 0% | (Not | 2 hours |

| Task | | | Assignee | | Priority | % | Status | Estimate |
|---|---|---|---|---|---|---|---|---|
| Introduction, Works Cited Page at end (dataset sources; copy & paste from ETL report)<br><br>DONE by Wednesday lunchtime | | | | Vander Hoek | | | started) | |
| Technical Report Research on Federal Questions (Section I)<br><br>First Section of technical report; include visualizations (screen shot from dashboard) for answering federal level questions (two questions) | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 3 hours |
| Technical Report Research Section II: Explain Machine Learning model, predictions, accuracy, include visualizations; answers questions 3/4 (predict poverty level using factors) | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 3 hours |
| Technical Report Research Section III: Answer questions about state (excluding machine learning); include visualizations | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 3 hours |
| Technical Report: Conclusion, Edit/Review other sections & introduction | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 2 hours |
| Class Presentation<br><br>Presentation for the class! | | | Anyone | Beth Vander Hoek | High | 0% | (Not started) | 20 minutes |
| All project components COMPLETE & PERFECTED  Blocker<br><br>BLOCKER FOR UPLOADING FINAL COMPONENTS TO GITHUB<br>EOD THURSDAY!!! Project is perfect :) | | | Anyone | Beth Vander Hoek | High | 0% | (Not started) | 2 hours |
| All PDFs, code uploaded to GitHub: presentation slides, final report, dashboard/visualizations code, ERD, napkin drawings & feedback (Dashboard & visualizations), ETL report, ReadMe, Azure Studio Links, data pipeline diagram, exploratory questions  Blocker<br><br>BLOCKER FOR FINAL README<br>THURSDAY MORNING/WEDNESDAY NIGHT ALL DONE!! | | | Anyone | Beth Vander Hoek | High | 0% | (Not started) | 3 hours |
| Update Repo/ReadMe as needed; make sure everything is uploaded properly<br><br>Thursday morning before presentation start | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 1 hour |
| Final Presentation<br><br>Final presentation to class, attendees, Dev10, Genesis10, shareholders, clients | | | Anyone | Beth Vander Hoek | High | 0% | (Not started) | 20 minutes |
| ETL Report Load Section | | | Anyone | Beth Vander Hoek | Medium | 0% | (Not started) | 1 hour |
| ETL Report Extract & transform: cpi  ● To do | | | Shannon Bayless | Beth Vander Hoek | | 0% | (Not started) | 0 |
| | | | Jerad Ipsen | Beth | | 0% | (Not | 0 |

| Task | | | | Responsible | Created By | | Progress | Status | Estimated |
|---|---|---|---|---|---|---|---|---|---|
| ETL Report Extract & transform: federal-party-control 🟡 In progress | | | | | Vander Hoek | | | (Not started) | |
| ETL Report Extract & transform: federal-spending 🔴 To do | | | | Jerad Ipsen | Beth Vander Hoek | | 0% | (Not started) | 0 |
| ETL Report Extract & transform: poverty-rates 🟡 In progress | | | | Shannon Bayless | Beth Vander Hoek | | 0% | (Not started) | 0 |
| ETL Report Extract & transform: state-crime 🟡 In progress | | | | Joel Garcia | Beth Vander Hoek | | 0% | (Not started) | 0 |
| ETL Report Extract & transform: state-finances 🔴 To do | | | | Shannon Bayless | Beth Vander Hoek | | 0% | (Not started) | 0 |
| ETL Report Extract & transform: state-od 🔴 To do | | | | Joel Garcia | Beth Vander Hoek | | 0% | (Not started) | 0 |
| ETL Report Extract & transform: state-party 🔴 To do | | | | Beth Vander Hoek | Beth Vander Hoek | | 0% | (Not started) | 0 |
| ETL Report Extract & transform: state-unemployment 🔴 To do | | | | Joel Garcia | Beth Vander Hoek | | 0% | (Not started) | 0 |
| ETL Report Extract & transform: us-poverty-rates 🔴 To do | | | | Shannon Bayless | Beth Vander Hoek | | 0% | (Not started) | 0 |
| | | | | | | | | | **61 hours 20 minutes** |

## Completed Tasks

| Task | Start Date | Due Date | Responsible | Created By | Priority | Progress | Status | Estimated |
|---|---|---|---|---|---|---|---|---|
| Build Producer and Topic 🟢 Completed<br><br>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 | 100% | Completed 09/21/2022 by Shannon Bayless | 2 hours |
| Create Repository 🟢 Completed<br><br>Create repo for capstone project; update as needed | | 09/16/2022 | Beth Vander Hoek | Beth Vander Hoek | | 100% | Completed 09/20/2022 by Beth Vander Hoek | 10 minutes |

| Task | Status | Start Date | Due Date | Assignee | Reporter | Priority | % | Completion | Time |
|------|--------|-----------|----------|----------|----------|----------|---|-----------|------|
| Brainstorm Questions; Search for Datasets; EDA ● Completed | | | 09/19/2022 | Dev10 | Beth Vander Hoek | | 100% | Completed 09/20/2022 by Beth Vander Hoek | 5 hours |
| Upload all data sets to azure container ● Completed | | 09/21/2022 | 09/21/2022 | Dev10 | Shannon Bayless | High | 100% | Completed 09/22/2022 by Jerad Ipsen | 30 minutes |
| Create ERD ● Completed | | 09/21/2022 | 09/21/2022 | Shannon Bayless | Shannon Bayless | Medium | 100% | Completed 09/22/2022 by Jerad Ipsen | 2 hours 30 minutes |
| Executive Summary - First Draft ● Completed | | 09/20/2022 | 09/20/2022 | Beth Vander Hoek | Shannon Bayless | Medium | 100% | Completed 09/20/2022 by Beth Vander Hoek | 1 hour |
| Napkin drawings data pipeline diagram `Blocker` ● Completed  How cloud/data factory flow will look like | | | 09/21/2022 | Joel Garcia | Beth Vander Hoek | Medium | 100% | Completed 09/21/2022 by Joel Garcia | 1 hour 30 minutes |
| Finish PMP ● Completed | | 09/20/2022 | 09/20/2022 | Beth Vander Hoek | Shannon Bayless | High | 100% | Completed 09/20/2022 by Beth Vander Hoek | 2 hours |
| Finalize Datasets ● Completed | | 09/20/2022 | 09/20/2022 | Anyone | Shannon Bayless | High | 100% | Completed 09/20/2022 by Beth Vander Hoek | 2 hours |
| Research Best Machine Learning Model for Data ● Completed  Research which ML model to use for data | | | 09/21/2022 | Jerad Ipsen | Beth Vander Hoek | | 100% | Completed 09/22/2022 by Jerad Ipsen | 2 hours |
| Create tables in Database based on ERD schema ● Completed | | | 09/22/2022 | Shannon Bayless | Beth Vander Hoek | | 100% | Completed 09/23/2022 by Beth Vander Hoek | 1 hour |
| Written Executive Summary (1 paragraph) ● Completed  Executive summary and list of questions | | 09/20/2022 | 09/20/2022 | Beth Vander Hoek | Beth Vander Hoek | Medium | 100% | Completed 09/20/2022 by Beth Vander Hoek | 2 hours |
| Build Consumer & Append to SQL database table ● Completed | | | 09/23/2022 | Shannon Bayless | Beth Vander Hoek | | 100% | Completed 09/23/2022 by Beth | 2 hours |

| | | | | | | | | | Vander Hoek | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 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 | | | | | | | | | |
| | Exploratory Questions pdf (5-10 finalized questions) 🟢 Completed | | | 09/20/2022 | Beth Vander Hoek | Beth Vander Hoek | Medium | 100% | Completed 09/20/2022 by Beth Vander Hoek | 1 hour |
| | Final data pipeline diagram 🟢 Completed<br><br>Maybe use icons to look pretty like GitHub repo project diabetes example? | | | 09/26/2022 | Joel Garcia | Beth Vander Hoek | Medium | 100% | Completed 09/23/2022 by Shannon Bayless | 3 hours |
| | Create two group Containers & SQL Database  Blocker  🟢 Completed<br><br>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 | 100% | Completed 09/20/2022 by Beth Vander Hoek | 1 hour |
| | Create Data Factory & Run Producer/Consumer Automated  🟢 Completed<br><br>Use exercise data factory as baseline. | | | 09/26/2022 | Beth Vander Hoek | Beth Vander Hoek | Medium | 100% | Completed 09/23/2022 by Beth Vander Hoek | 1 hour |
| | Add rest of cleaned data to SQL database (if not through producer/consumer data factory) 🟢 Completed<br><br>Make sure no repeats of data; only if not already added. Keep track of steps to add for ETL (merging, etc.) | | | 09/26/2022 | Shannon Bayless | Beth Vander Hoek | Medium | 100% | Completed 09/23/2022 by Beth Vander Hoek | 3 hours |
| | Clean & transform dataset: Cpi  Blocker  🟢 Completed | | | | Shannon Bayless | Beth Vander Hoek | High | 100% | Completed 09/23/2022 by Beth Vander Hoek | 2 hours |
| | Clean & transform dataset: federal-party-control  Blocker  🟢 Completed | | | | Jerad Ipsen | Beth Vander Hoek | High | 100% | Completed 09/23/2022 by Beth Vander Hoek | 2 hours |
| | Clean & transform dataset: federal-spending  Blocker  🟢 Completed | | | | Jerad Ipsen | Beth Vander Hoek | High | 100% | Completed 09/23/2022 by Beth Vander Hoek | 2 hours |
| | EDA: cpi  Blocker  🟢 Completed | | | | Shannon Bayless | Beth Vander Hoek | | 100% | Completed 09/23/2022 by Beth | 2 hours |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | Vander Hoek | |
| Clean & transform dataset: poverty-rates  Blocker  🟢 Completed | | | Shannon Bayless | Beth Vander Hoek | High | 100% | Completed 09/23/2022 by Beth Vander Hoek | 2 hours |
| EDA: federal-party-control  Blocker  🟢 Completed | | | Jerad Ipsen | Beth Vander Hoek | | 100% | Completed 09/23/2022 by Beth Vander Hoek | 2 hours |
| Clean & transform dataset: state-crime  Blocker  🟢 Completed | | | Joel Garcia | Beth Vander Hoek | High | 100% | Completed 09/23/2022 by Beth Vander Hoek | 2 hours |
| EDA: federal-spending  Blocker  🟢 Completed | | | Jerad Ipsen | Beth Vander Hoek | | 100% | Completed 09/23/2022 by Beth Vander Hoek | 2 hours |
| Clean & transform dataset: state-finances  Blocker  🟢 Completed | | | Shannon Bayless | Beth Vander Hoek | High | 100% | Completed 09/23/2022 by Beth Vander Hoek | 1 hour |
| EDA: poverty-rates  Blocker  🟢 Completed | | | Shannon Bayless | Beth Vander Hoek | | 100% | Completed 09/23/2022 by Beth Vander Hoek | 2 hours |
| Clean & transform dataset: state-od  Blocker  🟢 Completed | | | Joel Garcia | Beth Vander Hoek | High | 100% | Completed 09/23/2022 by Beth Vander Hoek | 2 hours |
| EDA: state-crime  Blocker  🟢 Completed | | | Joel Garcia | Beth Vander Hoek | | 100% | Completed 09/23/2022 by Beth Vander Hoek | 2 hours |
| Clean & transform dataset: state-party-control  Blocker  🟢 Completed | | | Beth Vander Hoek | Beth Vander Hoek | High | 100% | Completed 09/23/2022 by Beth Vander Hoek | 4 hours |
| EDA: state-finances  Blocker  🟢 Completed | | | Shannon Bayless | Beth Vander Hoek | | 100% | Completed 09/23/2022 by Beth | 2 hours |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | Vander Hoek |
| Clean & transform dataset: state-unemployment  Blocker  🟢 Completed | | | Joel Garcia | Beth Vander Hoek | High | 100% | Completed 09/23/2022 by Beth Vander Hoek | 2 hours |
| EDA: state-od  Blocker  🟢 Completed | | | Joel Garcia | Beth Vander Hoek | | 100% | Completed 09/23/2022 by Beth Vander Hoek | 2 hours |
| EDA: state-party  Blocker  🟢 Completed | | | Beth Vander Hoek | Beth Vander Hoek | | 100% | Completed 09/23/2022 by Beth Vander Hoek | 2 hours |
| EDA: state-unemployment  Blocker  🟢 Completed | | | Joel Garcia | Beth Vander Hoek | | 100% | Completed 09/23/2022 by Beth Vander Hoek | 2 hours |
| EDA: us-poverty-rates  Blocker  🟢 Completed | | | Shannon Bayless | Beth Vander Hoek | | 100% | Completed 09/23/2022 by Beth Vander Hoek | 2 hours |
| | | | | | | | | **71 hours 40 minutes** |