DS Capstone Project Guidelines

Get hyped. The last 3 weeks of the program are all about your Capstone project, which will showcase everything you've learned to employers. This is awesome!

Schedule

To help you make the most of the time, **we will begin preparing for Capstone at the beginning of Mod 4**. We will facilitate activities so you can hit the ground running on Day 1 of the Capstone weeks.

After Mod 4, your time is largely your own to work on your Capstone project. There will be a few lectures and other activities, but most of the time will be dedicated to your project. Think of this as like a real work environment, where you are responsible for managing your time and keeping yourself on-task.

Expectations

We reiterate the following expectations to help you make the most of our time together and leave with a portfolio-ready Capstone project.

- Show up by 9:00, leave no sooner than 6:00, and don't take extended breaks
- Participate in all stand-ups, check-ins and whiteboarding
- Work to meet all milestones, and communicate proactively about any issues or delays with your instructors
- Stick to the timeline and any decisions made with instructors
- After Mod 4, meet with your cohort for stand-up in the morning and stand-down at the end of the day
- Use the PEP: Research and debug before escalating to your coach/lead
- Be courteous to your classmates, and work together to solve problems

Stand-Up / Stand

When networking and in interviews, you'll find yourself saying a quick 30-second summation of your capstone (and the projects you'll do after bootcamp) **a lot**. To get practice succinctly describing your project in an "elevator pitch" format, there is a script you'll recite at stand-up and stand-down each day:

“Hello, my name is _____ and I’m a data scientist with a background in _____.”
Provide a 30-second description of problem / need / decision someone has to make that you’re addressing, and how your project addresses it. Be sure at some point to talk about what you predict / classify, and with what techniques.

Then, answer the following prompts:

1. **Feelings:** I’m feeling _____ about _____ because _____.
2. **Accomplishments:** Since we last spoke, I _____ .
3. **Plans for today:** Today I plan to _____ , the first step of which is _____ .
4. **Blocking issues (if needed):** I expect _____ to be a challenge / don’t currently know how I’ll tackle _____, and need some assistance.

Capstone Project

This is the time to dive into the data science process using techniques we’ve learned, and really put your skills on display via a self-directed project.

Because we’re asking you to show off a specific set of skills, we have some requirements. One of the requirements is that you need to implement the knowledge that you learned while at Flatiron. This isn’t the time to build a new classification algorithm or do reinforcement learning. You’ve done a ton of learning already - it’s time to **apply** all of that knowledge.

It is entirely up to you to select both the data you will use and question to answer. When narrowing down the options for your project, please refer to the Capstone Proposal Review Guidelines. This is intended to get you thinking about project management and how you are going to allocate your time.

Project Guidelines

Final project approval is up to your coaches and leads, who will be acting as project managers. You are **required to** get approval for your project, otherwise you will not be able to move forward on the project and therefore graduate.

Your final project is an elevated end-of-mod project. Your project must include some modeling. It can be from any of the topics that we've covered over the past 12 weeks and should demonstrate an understanding of the data science concepts you learned at Flatiron. Consider if your project will be relying on models or methods that we have covered so far, or going to be incorporating new ones. Try to keep the amount of new material to a minimum to ensure the last few weeks are spent on working on your project, rather than learning new concepts.

In the coming days, you will be asked for at least two potential ideas for your final project. Your coaches will help pick one idea, and then you will be responsible for writing up a proposal for lead approval. Once you receive approval, you can begin work on your project. Proposals will be submitted via Canvas.

Elements that Must be Included:

- A prediction on a new datapoint
- Input features and an output prediction
- Use of one of the models listed below
- Be able to be completed in less than 3 weeks

The dataset you ultimately select must contain at least 1000 observations.

Models to Choose From

Here are the models you can use. See section on special project permissions if the model you would like to use is outside of this list.

- Regression models (linear, CART, etc)
- Classification models (KNN, CART, logistic, etc)
- Tree Based models (decision tree, random forest, boosting)
- Time series models
- Neural Networks (limit to basic NN, CNN and RNN)
- Recommendation systems

Data Science Concepts and Tools

- Databases (SQL, MongoDB, etc)
- API Interaction
- Clustering
- Hadoop/Spark components
- Natural Language Processing
- Image Processing
- A webapp to showcase your project

Special Project Permission

If you would like to use a model outside of the above list you must notify instructors before your initial coach pitch. You will need to show proficiency in this model and provide evidence that you can utilize it for the purpose intended in your project early within the project proposal process. In addition, staff support may be extremely limited in assisting you with any roadblocks regarding your model.

Helpful Tools

Here are some of our favorite tools that may be helpful as you work on your Capstone project:

Kanban/Scrum Board

This will be the most complex project you've made at Flatiron, so you will want to create some structure to keep yourself organized and productive. We recommend Trello or a Github Project Board. Use this to track what you're doing and what you need to work on. It's also a great idea to keep track of bugs or issues that you're not going to immediately fix.

Pomodoro Timer

If you don't take breaks, you'll end up hurting your eyes, getting an RSI, or burning yourself out. The Pomodoro Timer method lets you put in solid chunks of work while also giving you regular breaks. We like Marinara Timer, since it's nicely customizable.

Other Tools and Services

Postman - Test API calls

Heroku - Simple, free web hosting for flask or django in python

DB Browser - SQLite interface for making calls to a database

AWS - Amazon Web Services for running models on the cloud