Phase 4 Expectations
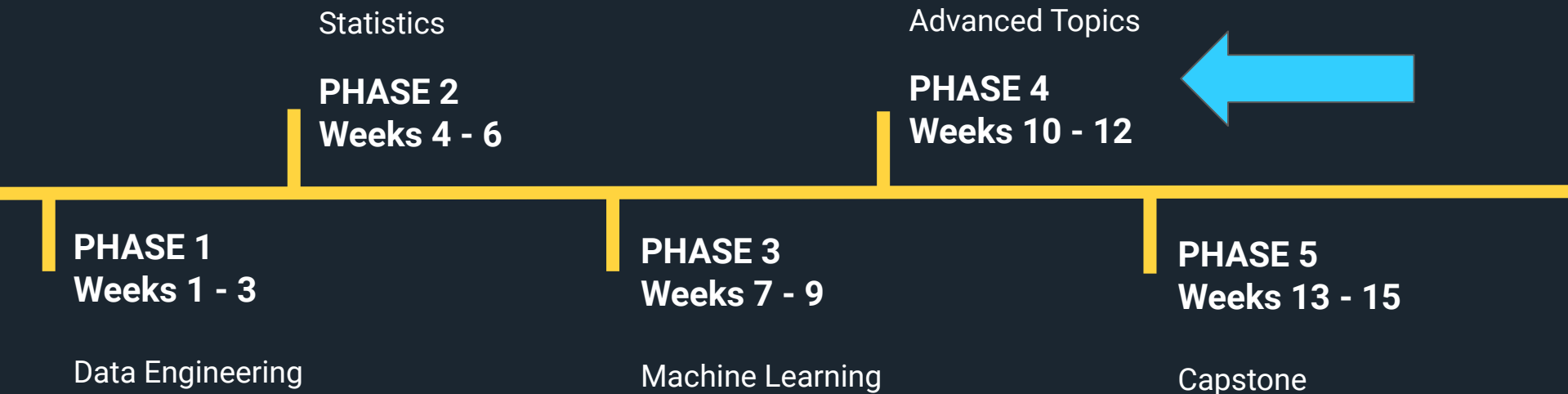
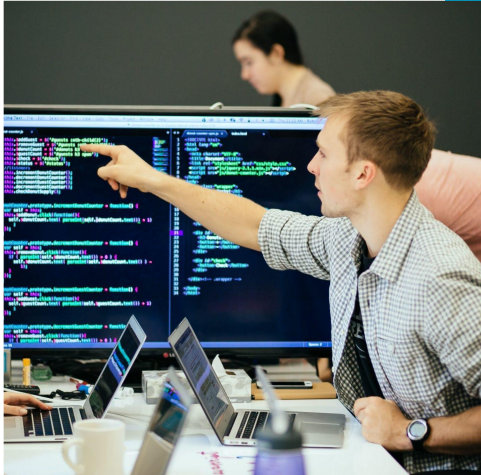# Agenda

- Timeline
  - Capstone coming up
- New tools/algorithms
- New data types
- Advance models

Overall Phase Timeline

Statistics

PHASE 2
Weeks 4 - 6

Advanced Topics

PHASE 4
Weeks 10 - 12

PHASE 1
Weeks 1 - 3

Data Engineering

PHASE 3
Weeks 7 - 9

Machine Learning

PHASE 5
Weeks 13 - 15

Capstone

// FLATIRON SCHOOL

Capstone Brainstorm

Answer the following questions for three (3) capstone ideas:

- ○ What real-world problem would your project aim to solve?
- ○ What data do you plan to use and what would be your target variable(s)
- ○ What methods would you use to model the data, and what would be the final product(s)?

- Due **Monday 01/15 EOD**

- Why three? So you have backup options!

- Schedule 1:1 Check Ins in Week 2 to connect and discuss

- Full Capstone proposal due Day 1 of Phase 5 (01/29)

Tasting Menu of New Topics

New Tools

Models

-K-Nearest Neighbor

-Random Forest

-Ensemble

-Boosting

Techniques

-Grid Searching

-Pipelines

-Clustering

-PCA

Data

-Text data (NLP)

-Ratings data

-Image data

Packages

-Surprise

-NLTK

-Tensorflow/Keras

Measuring Student Progress

Blog Post

Due 2nd **Tuesday** of Phase 4 (01/16)

-Summarize a data science research paper/project

-Write another technical tutorial

Checkpoints + Canvas

-Unsupervised ML Checkpoint: **Friday 01/12**

-NLP Checkpoint: **Tuesday 01/16**

-Tensorflow Checkpoint: **Thursday 01/18**

Code Challenge

Friday 01/19

-Pipelines

-Ensemble Models

-Unsupervised Machine Learning

-Natural Language Processing

Phase 4 Project

Options!

-Choose your own data

-Recommendation System

-Twitter Data

-Image Classification

Phase 4

Pick a research paper/case study about data science and write a blog post to explain it to a non-technical business stakeholder. This can be a new paper from the past few months, or you can refer to <u>this spreadsheet</u> for a list of historical papers. Potential elements to include:

I. What was the original context of this paper when it was written?
II. Summary of the paper findings/outcomes
III. How can this paper inform your work as a junior data scientist?
IV. Why is this paper important/why does it matter to a non-technical business stakeholder?

Feelings at the Start of Phase 4

Confident about what's going on...kind of

Frustrated by ongoing "beginner" mistakes

Overwhelmed by studying new concepts while fleshing out capstone.

Feelings at the End of Phase 4

Cautiously optimistic when things start to "click"

Struck by the reality of the bootcamp ending

Prepared to get started on capstone

LET'S DO THIS.