Jeremy the Junior

Age: 18 - 21
Occupation: Undergraduate Student

Scenario

Student in STATS 68 (Statistical Computing and Exploratory Data Analysis) course. Professor assigns a collaborative project that requires use of JupyterLab. Jeremy is an occasional user of JupyterLab. Professor uploads data to cloud storage and expects students to access and use it.

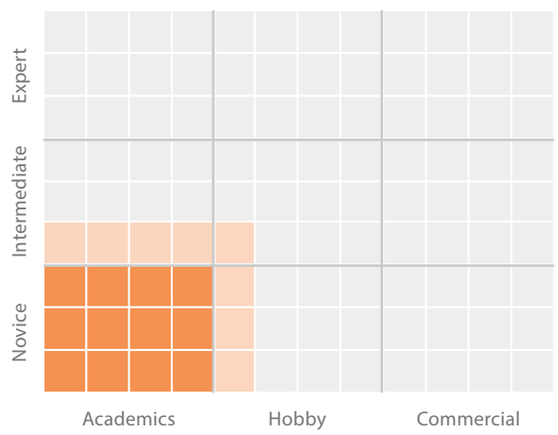
Quotes

“I wish I could have easier access to the files.”

“I didn’t understand that what I see in JupyterLab is actually my file.”

Motivations

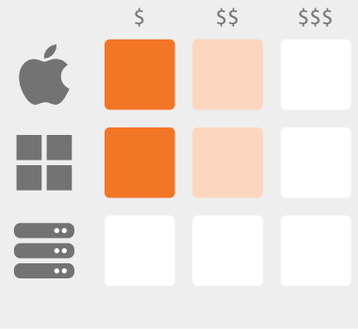
Expertise Matrix

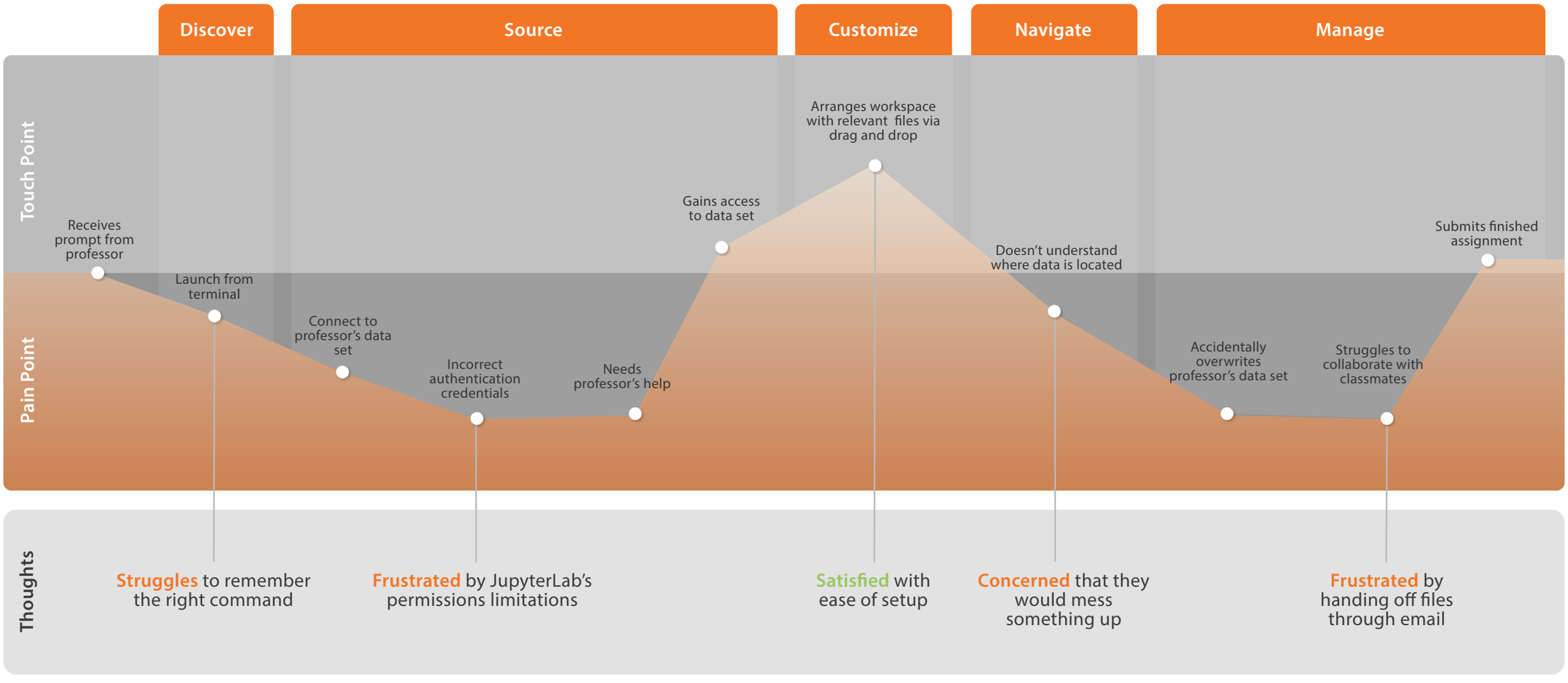


Goals and Expectations

- Learn data science with instructor-provided data sets
- Group projects with peers that follow non-linear working patterns are common
- Expects most data to be locally stored and readily accessible

Computing Power





Isaac the Intern

Age: 24
Occupation: Intern

Scenario

An intern at a health science research lab. Isaac supports the senior level researchers who are exploring viral mutation rates. He is working together with a senior researcher to create data models for a published paper.

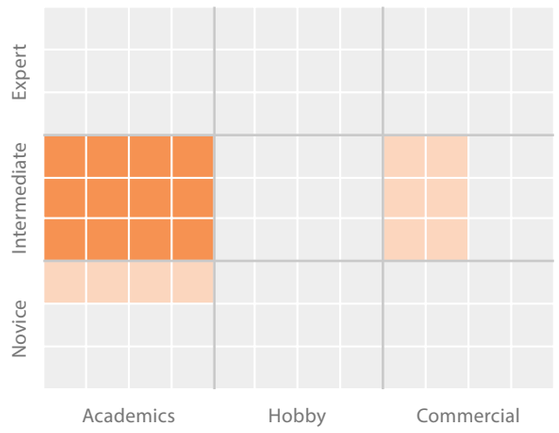
Quotes

“I wish it were easier for me to rename new files.”

“It would be nice to have easier access to remote files.”

Motivations

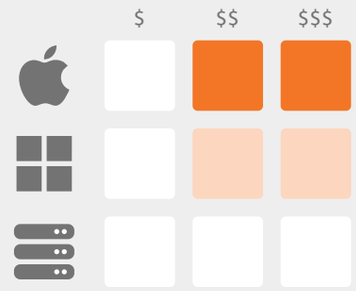
Expertise Matrix

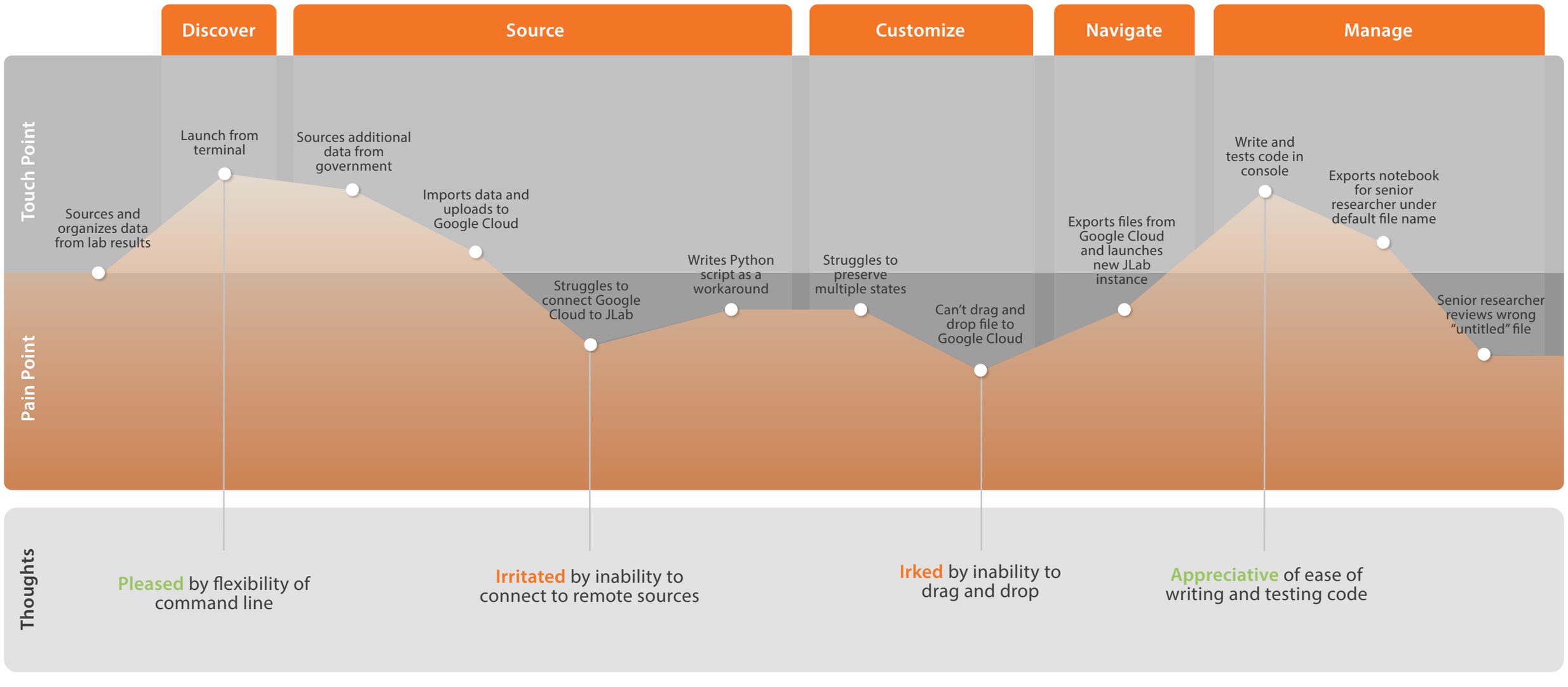


Goals and Expectations

- Establish reputation in the data science community
- Wants to view data from different lenses to ensure the validity of his models
- Uses diverse data sources in a cohesive fashion
- Produce work that can be replicated

Computing Power





Denise the Developer

Age: 32
Occupation: Developer

Scenario

Developer at an investment firm who works daily with a portfolio management application. She uses public and private data sets to create investment models to project market trends.

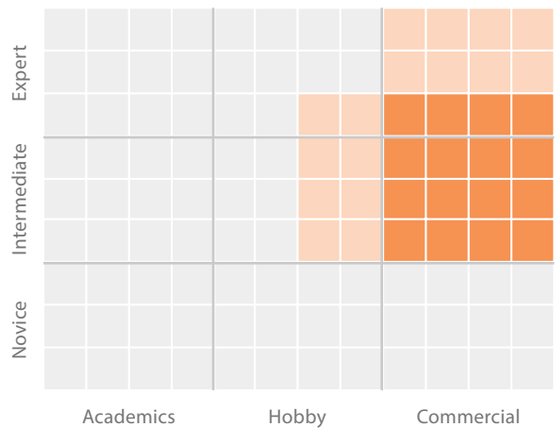
Quotes

"I wish I could preview large data sets in JupyterLab."

"I'm keeping track of a lot of different states at the same time."

Motivations

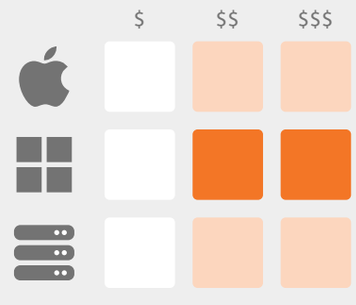
Expertise Matrix

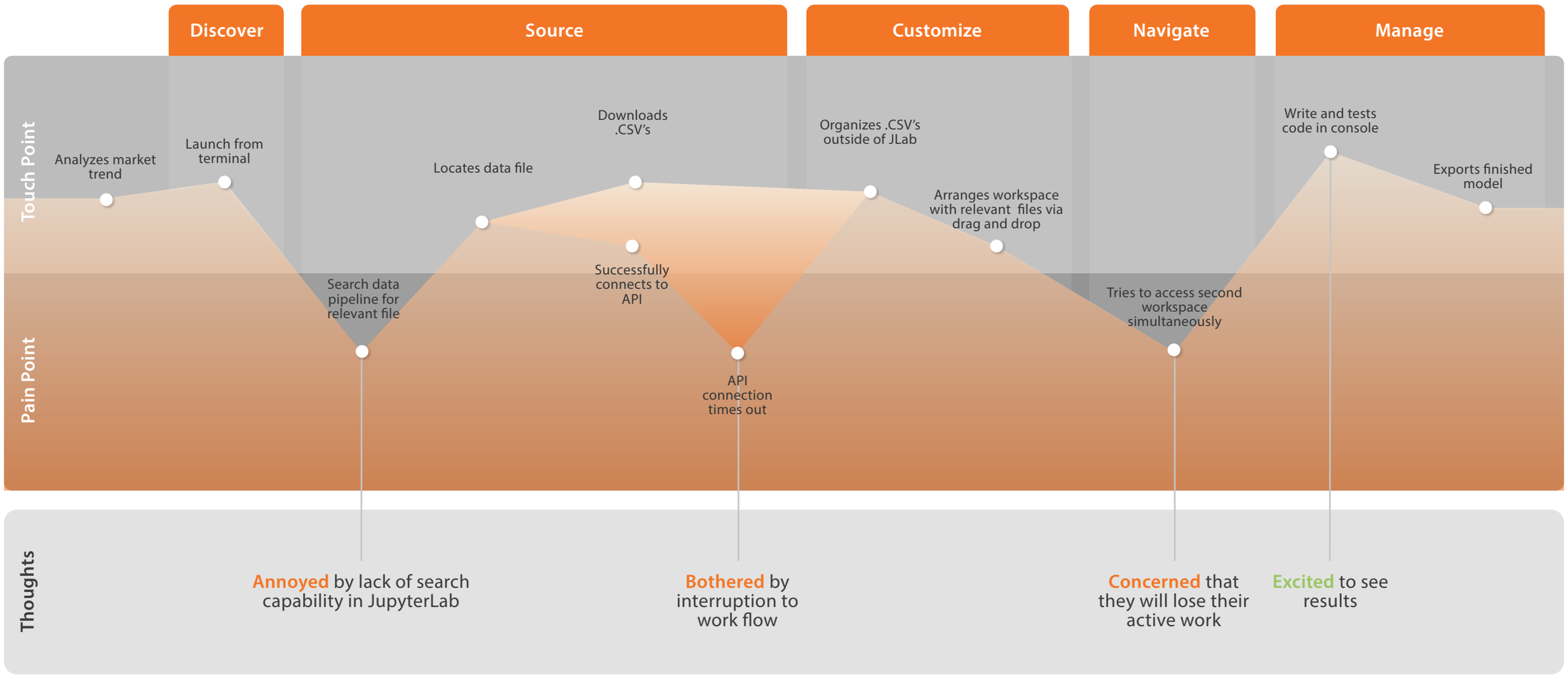


Goals and Expectations

- Expects to receive data from multiple sources outside of JupyterLab
- When working with collaborators, expects seamless handoff
- Expects to continuously update data sets and re-run investment models

Computing Power





Salma the Scientist

Age: 46
Occupation: Sr. Data Scientist

Scenario

Senior data scientist who collaborates with peers often and sources her data daily when using JupyterLab. She is working on a report for senior leadership regarding recent customer trends.

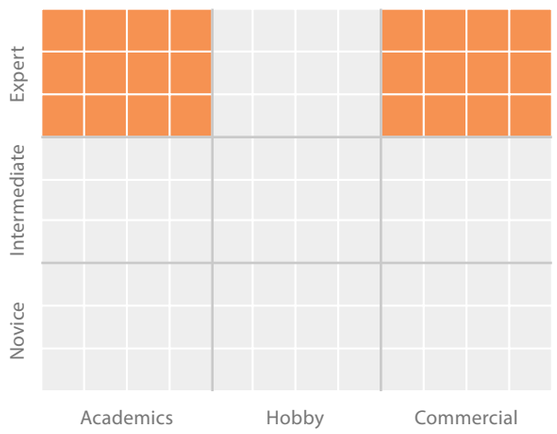
Quotes

"We data scientists could live on .CSV files."

"If you could pop open the first five rows of a .CSV file in a matter of milliseconds, that would be great."

Motivations

Expertise Matrix



Goals and Expectations

- Maintain data pipeline
- Ability to "free the data" from its host file system structure
- File versioning/history
- Produce cutting-edge data science models
- Ability to easily work with wide variety of data types

Computing Power