Cognitiv Take Home Challenge

This take home can be written in any JVM-based language, the only ask is that the result is an archive that contains the source code, test data and any needed build scripts so we can run it ourselves. Feel free to use outside libraries, frameworks, etc. as needed. Besides a JVM, there should be no other system dependencies to build or run your solution (e.g. no external databases, web services). The solution should run as a command-line process.

Included in your project should also be a `README.md` with instructions on how to set up and run your project, as well as any tests that you've written. The project can be packaged up and sent as a tarball.

Things to focus on while building this project:

1. **Readability.** The code should be well-structured and easily understood (please include comments like you would writing production code)
2. **Performance**. This dataset is not large enough to require anything particularly exotic in terms of data structures or algorithms, but bonus points for explaining how you might scale or improve the performance of your solution.
3. **Ease of setup**. Should be able to unarchive, build and run with minimal effort. Please document any needed prerequisites such as the JVM version, command-line arguments, and other necessary setup information.

Dataset:

Below is a link to NYC taxi data from the last few years:
https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page
(Also on the page are data dictionaries for the various bits of data).

The problem:

Please build a command line application that takes the last 2 months of data (include them in the project). We would like to be able to find aggregate information (min fare, max fare, count, sum of fares, and sum of toll fares, counts by payment type) of trips over the following dimensions:

1. Trips between certain times
2. Trips between specific locations
3. Trips from a certain vendor
4. Trips for types of taxis (yellow, green or both)

These dimensions should be able to be combined (i.e I'd like to know trips from a vendor over a certain time between these locations).