**Project: Watching the Stock Market**

This project will take you off-platform and get you started in your own developer environment! Never done that before? Not to worry - we’ve shared some resources to help you down below. This project can be completed entirely on your own - or, you can join the #data-science-buddies [in the Codecademy Pro Learner community 45](https://codecademy.com/users/redirect?redirect_url=https://join.slack.com/t/codecademy-school/shared_invite/enQtNDUzNjc0NTA3OTI3LTBkM2YwYTAxMzM0OWVmNjc1YjY5MjNhYjc5OWE0MjgwZTliOGViNzhkNzhlZWExYTA3YjU3MDgyNGY0N2FhYTM) on Slack and find someone to work with! Jump to the community support section to hear more about this.

This project is broken down into key questions that your client or company is looking to answer. As a data scientist, you’ll often become a resource to help businesses answer the key questions about the efficacy of existing or potential strategies & projects.

**Overview**

**Objective**

You are asked by a company to help them make more informed decisions on investments. To start, you will be watching the stock market, collecting data, and identifying trends!

**Pre-requisites**

In order to complete this project, we suggest that you have familiarity with the content in the following courses or lessons on the Codecademy platform:

1. [What is a Relational Database Management System? 42](https://www.codecademy.com/paths/data-science/tracks/dspath-why-data-python-basics/modules/dspath-introduction-to-sql/articles/what-is-rdbms-sql)
2. [Manipulation 38](https://www.codecademy.com/paths/data-science/tracks/dspath-why-data-python-basics/modules/dspath-introduction-to-sql/lessons/manipulation)
3. [Queries 20](https://www.codecademy.com/paths/data-science/tracks/sql-basics/modules/dspath-sql-queries/lessons/queries)

**Suggested Technologies**

Depending on where you are on your Path, there may be multiple technology options you can use to complete this project - we suggest the following:

1. [DB Browser for SQLite 168](https://sqlitebrowser.org/)

**Project Tasks**

**Get started - hosting your project**

[DB Browser for SQLite 168](https://sqlitebrowser.org/) is a visual tool for working with SQLite databases. Follow the link to download the application for your computer.

https://lh3.googleusercontent.com/a-rAsiOjlukNezJgcI3gTowF9sFsvU3sEIZ-QiufSU-bABod6S1mz9D1yJ9T9LXMv0KVEj0ACo7VHReJwd_1ZFpidguqGIGtnHId8-7J516dMWjAvGiyWmsDs8husYnqXEKZmcHR

* SQLite can store an entire database in a single file, which usually has a .sqlite or .db extension. To open a database, select Open Database at the top of the window and browse for the file. Alternatively, you can choose to create a New Database by saving a file with the .sqlite or .dbextension.
* To import data from a CSV file into a table, select “File > Import > Table from CSV file…” and browse for the CSV file. (Note: All fields imported from the CSV file will have a data type of TEXT. Be sure to convert fields to numeric type as needed. See [here 21](https://stackoverflow.com/a/20241031/6373540) for how to do that.)

https://lh4.googleusercontent.com/gk3-6NlpRhpKHefZUa2Te1ZTSkgfK6p79u2Qaq-Ys8viQYzDV-IPVe8DXi-uy8BAII20JxzZUhRliFya17JxW8x3z-HQ3u_w9iKOxfKkLQPOrqssY3QtXp_rj-wO7fpEro2v44OB

There are several tabs near the top of the window for working with the data:

* Database Structure: View the tables in your database and the columns they contain.
* Browse Data: Browse the data for each table.
* Execute SQL: Write and execute SQL queries.

**Basic Requirements**

Let’s break this project down into a couple different parts.

Manipulation: Collect data on your pick of 5 [stocks 86](https://www.cnbc.com/stocks/).

* Create a table called stocks, where you will be inserting your data.
  + Hint: See [here 46](https://www.w3schools.com/sql/sql_create_table.asp) for a review of the CREATE TABLE syntax. What [data type 31](https://www.tutorialspoint.com/sqlite/sqlite_data_types.htm) should each field be?
* The stocks table should have a column for symbol, name, datetime, and price.
* Collect your data! Choose 3 times throughout the day to document the price of each stock and continue for at least 1 week. You can do this moving forward, or just take a retroactive look at the stock market ﻿﻿by taking data historically from regular intervals (e.g. the first of the month for the last six months).
  + Hint: See [here 31](https://www.w3schools.com/sql/sql_insert.asp) for a review of the INSERT INTO syntax. When inserting the datetime, use the standard format ‘yyyy-mm-dd hh:mm:ss’. Use the [strftime() 28](https://www.w3resource.com/sqlite/sqlite-strftime.php) function to help you get the datetime of ‘now’.

Queries: Perform basic analysis on the data and identify trends.

* What are the distinct stocks in the table?
* Query all data for a single stock. Do you notice any overall trends?
* Which rows have a price above 100? between 40 to 50, etc?
* Sort the table by price. What are the minimum and maximum prices?

**Additional Challenges**

Intermediate Challenge

* Explore using aggregate functions to look at key statistics about the data (e.g., min, max, average).
* Group the data by stock and repeat. How do the stocks compare to each other?
* Group the data by day or hour of day. Does day of week or time of day impact prices?
* Which of the rows have a price greater than the average of all prices in the dataset?

Advanced Challenge

* In addition to the built-in aggregate functions, explore ways to calculate other key statistics about the data, such as the median or variance.
  + Hint: See [here 8](https://stackoverflow.com/a/15766121/6373540) and [here 6](https://stackoverflow.com/a/4621268/6373540) for possible solutions.
* Let’s refactor the data into 2 tables - stock\_info to store general info about the stock itself (ie. symbol, name) and stock\_prices to store the collected data on price (ie. symbol, datetime, price).
  + Hint: You can use the SQL CREATE TABLE AS statement to create a table by copying the columns of an existing table. Don’t forget to also drop certain columns from the original table and rename it.
* Now, we do not need to repeat both symbol and name for each row of price data. Instead, join the 2 tables in order to view more information on the stock with each row of price.
* Add more variables to the stock\_info table and update the data (e.g., sector, industry, etc).

**Resources & Support**

**Project-specific resources**

1. [SQLite Data Types 31](https://www.tutorialspoint.com/sqlite/sqlite_data_types.htm)
2. [SQLite Date and Time Functions 7](https://www.sqlite.org/lang_datefunc.html)
3. [SQLite strftime() Function 28](https://www.w3resource.com/sqlite/sqlite-strftime.php)
4. [SQLite Documentation 3](https://www.sqlite.org/doclist.html)
5. [SQLite Tutorial 10](http://www.sqlitetutorial.net/)

**General Resources**

1. [How to get set-up for coding on your computer 14](https://discuss.codecademy.com/t/guide-set-up-coding-on-your-own-computer/395168/2)
2. [What is a Relational Database Management System? 1](https://www.codecademy.com/articles/what-is-rdbms-sql)
3. [What you need to know about Git, GitHub & Coding in Teams](https://discuss.codecademy.com/t/what-you-need-to-know-about-git-github-and-coding-in-teams-guide-reading-list/394581)
4. [How developer teams work 2](https://discuss.codecademy.com/t/guide-how-developer-teams-work/394900)
5. [First steps in tackling a group project 1](https://discuss.codecademy.com/t/guide-first-steps-in-tackling-a-group-project/394959)
6. Resource on [writing pseudocode](https://docs.google.com/document/d/1MAOhbcP5JjId7ScflL6ZgIsft5Qi7m3WtaYV2Apu15c/edit?usp=sharing) to get started with off-platform projects

**Community Support**

Want to Looking for additional help or someone to work with (or somewhere to brag about your finished project)? Join our Codecademy Pro Learner Community on Slack to meet other learners like yourself!

* Collaborate with other learners on data collection! Then, join the datasets together for more interesting analysis.
  + Each learner can collect data on different stocks for a larger sample of stocks.
  + Each learner can collect data on same 5 stocks, but at different points throughout the day in order to spot potential daily trends.

**Once you’re done…**

Share on [Slack](https://www.codecademy.com/pro/community) or the [Forums 1](https://discuss.codecademy.com/c/project-feedback) for feedback and to see some other ways of solving this problem!