# Reading: Course 2 resources and citations

## Module 1: Data models and pipelines

### Resources

**Guide to Dataflow**

- Dataflow Google Cloud

**Python applications and resources**

- The Python Software Foundation (PSF)
- Python Tutorial
- Coding Club Python Tutorials

**Case study: Wayfair - Working with stakeholders to create a pipeline**

- Wayfair

**[Optional] Review Google Data Analytics Certificate content about SQL best practices**

- Sublime Text
- Atom
- Search and replace in Sublime Text
- Regex tutorial
- Regex cheat sheet

### Citations

**Python applications and resources**


- Coding Club tutorials. (n.d.). GitHub. https://ourcodingclub.github.io/tutorials.html
- Python for beginners. (n.d.). Python. https://www.python.org/about/gettingstarted/
- The Python tutorial. (2022, July 7). Python. https://docs.python.org/3/tutorial/


## Module 2: Dynamic database design

### Resources

**Indexes, partitions, and other ways to optimize**

- Devart's article on SQL Query Optimization

**Case study: Deloitte - Optimizing outdated database systems**

- Deloitte

## Citations

**Indexes, partitions, and other ways to optimize**

- SQL query optimization: How to tune performance of SQL queries. (2021, Dec 23). Devart. https://blog.devart.com/how-to-optimize-sql-query.html

# Module 4: Course 2 end-of-course project

## Resources

**Cyclistic datasets**

- NYC Citi Bike Trips, Census Bureau US Boundaries
- GSOD from the National Oceanic and Atmospheric Administration
- zip code spreadsheet
- the console

**Observe the Cyclistic team in action**

- NYC zip codes

**Google Fiber datasets**

- Market_1
- Market_2
- Market_3
- the console

**Activity Exemplar: Create your target table for Google Fiber**

- Create a target table for Google Fiber exemplar

---