DTSC 1301/2 Project Specifications

Fall 2022

Important Dates

- Section drafts due Tuesday, September 13; Thursday, September 22; and Thursday, September 29 **(all at 7pm)**
- Final written report must be submitted on Gradescope by **12pm (noon)** on October 12

Purpose and Goals

This project is a culminating experience for DTSC 1301+DTSC 1302. A "data science project" is a unit of work for many working data science professionals. In your first data science project, you will:

1. complete your first data science project from beginning to end.
2. use social science and humanities skills to clearly communicate the context of your project and its results.
3. develop technical and non-technical skills needed by data science professionals.
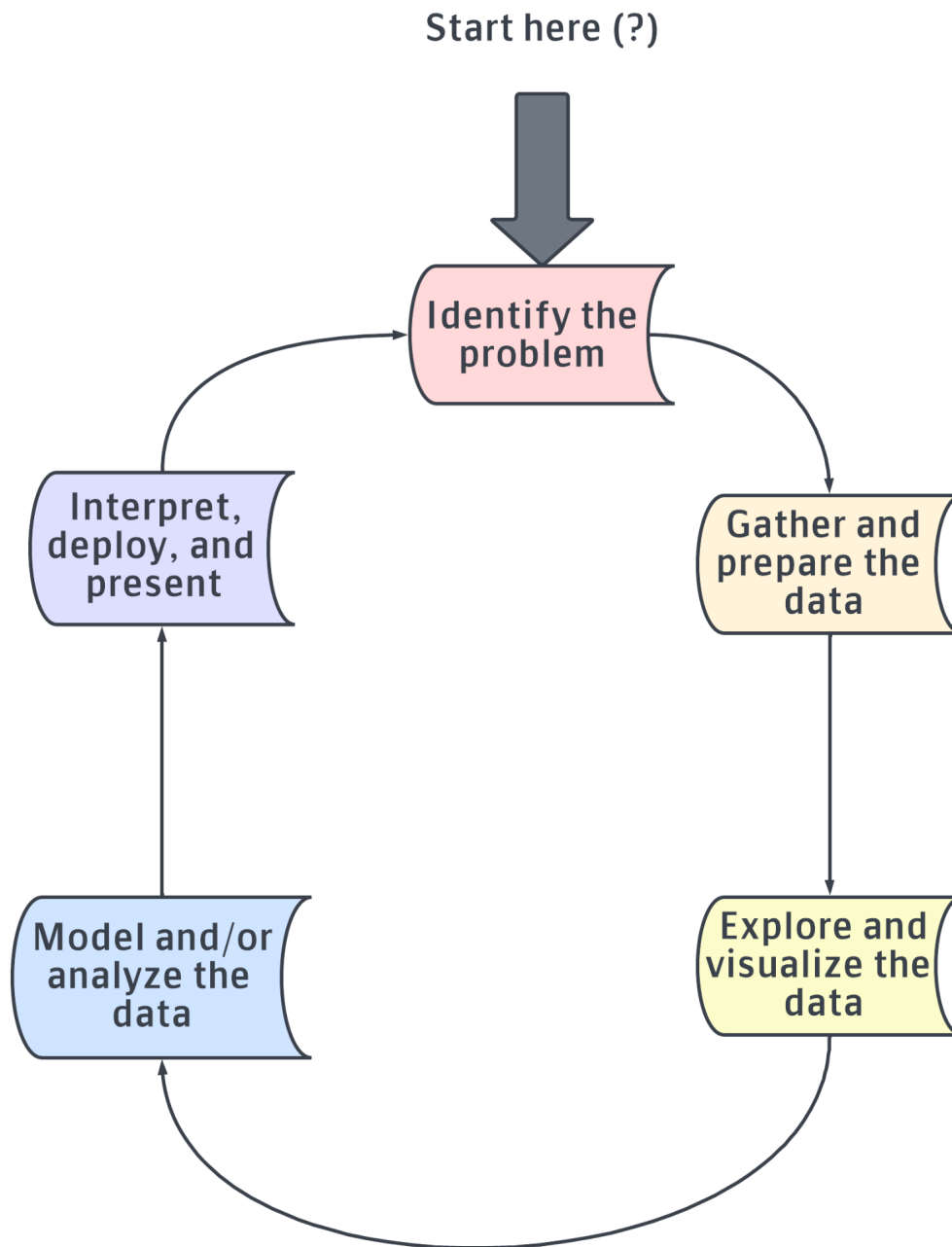
Start here (?)

Figure: The Data Science Life Cycle

Project Scenario

Charlotte is growing! [Sources: 1 2 3 4 and so many more!]

Your team is expected to describe clearly and precisely:
1. in what way(s) Charlotte is growing,

2. the rate of growth over a recent time period, and
3. what the implications of this period of growth are for stakeholders.

Your Task

Over the course of this semester, your team will:
- Formulate a research question targeting the items in the project scenario.
- Develop a theory around your research question, with supporting evidence.
- Articulate the context and implications of the question you've chosen to explore.
- Explore data relevant to your research question.
- Describe and visualize preliminary observations from your data exploration and relate these observations back to your research question, theory, context, and implications.

Part 1 Requirements

During the first half-term, your team will prepare a preliminary report. It must include the following four sections, in this order:
1. Introduction
2. Theory
3. Context and Implications
4. Works Cited

Part 2 Requirements

During the second half-term, your team will refine your preliminary report and prepare a final exploratory report. It must include the following six sections, in this order:
1. Introduction
2. Theory
3. Context and Implications
4. Data
5. Conclusion
6. Annotated Bibliography

Your team will submit all data files and Python files used to summarize and visualize the data to write and support your Data and Conclusion sections.

Your team will give an oral presentation during class time to describe your research question and its context and implications.

Section Details

Introduction

Your report introduction must clearly state your research question, and motivate why this question is important to explore.

Theory

Your report theory section must thoroughly explain the following:
- **Conceptualization**: All concepts used in your study, and how your research team is conceptualizing them.
- **Theoretical framework:** Use existing work (course readings and other published, peer-reviewed articles) to explain why it is reasonable that all of your concepts come together in your stated research question.
    - **Part 1:** You are not required to incorporate any sources above and beyond the assigned course readings.
    - **Part 2:** Your team will be expected to find additional appropriate sources to fine tune your theory.
- **Operationalization:**
    - **Part 1:** Describe how, in an ideal or perfectly designed study, your research team would operationalize each concept involved in your exploration of your research question.
    - **Part 2:** Describe how, using the data available to you for your project, your research team has operationalized each concept involved in your exploration of your research question.
- **Measurement**:
    - **Part 1:** Describe how, in an ideal or perfectly designed study, your research team would measure each concept involved in your exploration of your research question.
    - **Part 2:** Describe how, using the data available to you for your project, your research team has measured each concept involved in your exploration of your research question.

Context and Implications

You should describe the context of your research as completely as you can. In particular, you are expected to include:
- Historical conditions and events in the Charlotte region that are relevant to the specific populations impacted by your research question
- A description of all stakeholders in the question you chose to explore, their differing interests, and how they may be affected

- ○ Be sure to address any conflicting stakeholder interests!
- Any ethically significant benefits or harms to the stakeholders that may arise due to the themes explored in your research
- Describe your ethical approach in terms of how the referenced citations on data ethics, justice, and fairness shape your own data and contextual analysis

Data

In this course, your team will be provided with a template for writing this section of your report. This template will have space for you to describe the data set(s) used, describe all variables used and how they were measured, and summary statistics and visualizations for each variable used.

Conclusion

Your conclusion section should answer the following questions. However, it should be formatted as a report (using complete sentences and paragraphs), not as a bulleted list or FAQ-style Q&A.
- What observations can you make from the data that are relevant to your research question?
- What would the next steps be to answer your research question more rigorously?

Works Cited / Annotated Bibliography

Both of these sections involve writing a bibliography for your report. The bibliography should conform to one of the following citation formats: ACM, MLA, APA, or Chicago. The citation format must be consistent throughout your report, including in-line citations and references.

The *differences* between the expectations for part 1 and part 2 are described below:
- **Works Cited (part 1)**: A bibliography listing all sources consulted when doing part 1 of the project.
  - ○ You are not required to incorporate any sources above and beyond the assigned course readings.
- **Annotated Bibliography:** A bibliography listing all sources consulted when doing parts 1 and 2 of the project. Underneath each individual source reference, you must include 1-2 paragraphs explaining what the source is and why it is relevant to your work.
  - ○ Your team will be expected to find additional appropriate sources to fine tune your theory. The minimum number of additional sources is the number of students in your research team.