Updated Version.

CS 218 - Homework 2 Summary

Note that there are two due dates:

Part 1: Due by Midnight, Thu. Aug. 28

Create a google doc to answer the questions below (except team norms)

- Task 1.1: Reflecting on Teamwork and Class Norms ← Do it, but turn it in later with the other google docs parts below
- Task 1.2: Synthesize Class Norms ← Only this part is due by midnight, Thur, Aug 28

Use this this Form to submit **Class Norms**

Part 2: Due By 3:00pm, Wed. Sept. 3 (just before next class)

- Task 2: Spreadsheet Formulas and Charts
- Task 3: Correlations (7-Lens Data Scan Continued)
- Task 4: Prioritizing Neighborhoods for Housing Interventions
- Optional (Advanced)

Once you have answered all the questions in Tasks 1-3, use the
Homework Turnin form
to submit the google doc you created.
(Hint: Make screenshots small, since your document is limited in size to 10 MB.)

Instructions

This assignment is designed to simulate a realistic workplace challenge, where not every instruction is spelled out and you'll need to make thoughtful decisions along the way. **That ambiguity is intentional.** In many real-world data jobs, you'll be expected to figure things out, make reasonable choices, and communicate your reasoning.

Supports at your disposal:

- You've already seen tools and techniques in class that will help.
- You're encouraged—and expected—to use **ChatGPT** or similar tools to:
    - Help write or debug formulas
    - Ask about spreadsheet features or visualization methods
    - Explore and analyze data
- You can look up **tutorials online**
- You can ask questions in class, **consult with the TAs**, or collaborate with classmates.

We've built **rails and structure** into the assignment to keep it manageable, but we are deliberately leaving some tasks open-ended to foster **independent reasoning**. Your responses should reflect your **own thought process**—which can absolutely be shaped, supported, or even challenged by an AI tool.

TL;DR: Use all the tools you have. Think for yourself. Show your work.

Task 1.1 – Reflecting on Teamwork and Class Norms

Start by creating a google doc to store your answers for all the questions in Homework 2 below. (The exception is the "Propose Class Norms" section, which has a separate form.)

Step 1: Reflection
Read the following class materials and answer each of the questions below:

- CS 218 - Working in Teams Doc (each below is a tab in the document)
  - **Team Roles**
  - **Constructive vs. Destructive Behaviors**
  - **Team Meeting Types**

Then answer:

- Which team role do you think you're best at? Which is a growth area?
- Which constructive behavior comes naturally to you?
- Which destructive behavior do you tend to slip into—and how are you working on it?

Step 2: Propose Class Norms **(Only this Class Norms form is due by Midnight, Thursday, Aug 28). Everything else is due by 3pm Wed Sept 3.**
Come up with **3–5 norms** that you believe our class should follow to support good teamwork and culture.
Submit your norms via the Google form linked above in this document here, by **midnight, Thursday, Aug 28.**
You will need to be signed into your browser using your UIC identity to access this form.

Some classic examples of working norms might include things like:

- *Assume positive intent*
- *Use "yes and…" responses*
- *Always cite help from AI or classmates*

Task 1.2 – Synthesize Class Norms

- Find the spreadsheet with everyone's responses to part 1.1
  After the deadline (Thurs midnight), this will be listed at the bottom of the HW 2 page on the course website.

- Use ChatGPT (or another AI tool) to explore the responses, analyze, summarize and **rank a Top 10 Norms list**.
  If you are new to using ChatGPT for a job like this, then **watch** **this 4 min. video** that describes much of what you need to know to do this homework.
- You can rank by consensus, importance, or another method—just explain your reasoning.

Task 2 – Spreadsheet Formulas and Charts

Add your answers to the questions below to your google doc for homework 2. Use screenshots pasted into your google doc to show the results of your spreadsheet explorations.

If you are new to spreadsheets, then **watch** **this 35 min. video** that describes much of what you need to know to do this homework.

Scenario:

Your team's manager wants you to categorize **crowded housing** into **low/medium/high** problem areas and explore what that means for community analysis.

Do the following:

1. **Create a column** in the spreadsheet to classify each neighborhood as Low, Medium, or High crowded housing.
   - Use `IF()` (or nested `IF()` formulas!).
2. Make a mini table that uses **`COUNTIF()`** to show how many neighborhoods fall into each category of low/medium/high.
3. **Create a bar chart** of that table to show the distribution of low, medium, high.
4. Add a **new column** converting "Low/Med/High" into **numbers** 1, 2, 3.
5. **Compute the average** of these numbers.
6. **Reflect**: What does this average tell you? Does it accurately summarize the data in your opinion? What are the potential benefits and drawbacks of converting numerical data into categories/labels like low/med/high, and then giving those labels numbers (1,2,3) for further analysis?

(Commentary: a common pattern is to categorize numerical data into text labels, and then convert those text labels into numbers for

---

Task 3 – Correlations

Add your answers to the questions below to your google doc for homework 2. Use screenshots pasted into your google doc to show the results of your spreadsheet explorations.

1. Pick **one other field** in the Chicago Community Data you think might *correlate* with crowded housing. (We used median income in class, which you may use here. But you may pick another one of interest for a challenge or for fun.)
2. **Make a scatterplot** of the two variables.
3. **Label axes and title** the chart clearly.
4. Below the chart, write:
    - What pattern do you see?
    - Why do you think these two variables might be related?

Next:

5. Categorize your chosen variable into 3 labeled categories (You can use Low/Medium/High again if appropriate, or pick three different labels that make more sense).
6. Create a bar chart similar to the one you made in Part 2.

Once you are done creating the google doc with your answers to Tasks 1, 2, and 3 described above, turn in your google doc using the Homework Turnin form linked above in this document.

Task 4 shown below has been removed from this homework and will instead be part of the next homework.

Task 4 – Prioritizing Neighborhoods

Context:

A senior policymaker in your organization is preparing a citywide presentation to the mayor and council members. She wants to understand **which neighborhoods are in the greatest need of housing intervention.** She's already on board with using **crowded housing** as one of the key indicators — but she wants to **combine it with one other factor** to get a clearer picture of urgency.

> "I need a simple way to show where housing stress is worst. But keep it *simple*. I want a grid, or maybe a map - whatever makes the most sense - that shows how areas fall into categories. Make it something I can show and explain in 30 seconds. Don't give me scatterplots with regression lines — I need to *say something actionable*."

Add your answers to the questions below to your google doc for homework 2. Use screenshots pasted into your google doc to show the results of your spreadsheet explorations.

Your task:

1. **Identify** neighborhoods that are "High" for crowded housing **and** an extreme (High or Low) for another factor of your choice. For example, you might identify neighborhoods that are high for crowded housing and low for median income, or high for crime.
2. Create a **prioritized list** of 5-10 neighborhoods to focus on.
3. Draft a 30-second pitch for the policymaker: Why these neighborhoods?
4. Include a visual (filtered spreadsheet, bar chart, map, etc.) to aid the pitch.

Optional (Above and Beyond): Make a Metric

Frequently, organizations will want to develop a *metric*. A **metric** is a single number that condenses complex information into a clear signal, where everyone agrees whether "up" or

“down” is good. Its power is that it drives action — when the number changes, people know they should respond (even if they might not know *how* to respond). Classic examples of metrics are:

Finance & Economy

1. Dow Jones Industrial Average (DJIA)
2. NASDAQ Composite Index
3. Credit Score (FICO, etc.)

Sports

4. Batting Average (Baseball)
5. Quarterback Rating (NFL Passer Rating)
6. Player Efficiency Rating (PER, Basketball)

Business & Organizations

7. Net Promoter Score (NPS)
8. Employee Engagement Index

Everyday Life & Society

9. Body Mass Index (BMI)
10. Air Quality Index (AQI)

Add your answers to the questions below to your google doc for homework 2. Use screenshots pasted into your google doc to show the results of your spreadsheet explorations.

Stretch challenge (desired, but not required): Turn both Chicago community indicators into numbers and create a simple score (e.g., average or weighted sum). For example: develop a “crowded housing urgency” metric.

- Describe/define the metric.
- Use the metric to prioritize neighborhoods.
- How would you pitch this metric to your manager? How does it improve over “naive” methods of prioritizing neighborhoods?

Original Version [Deprecated]

CS 218 - Homework 2

Please note this is effectively two homework assignments because we do not have class on Monday. Therefore, there are two due dates to keep track of.

Part 1 - Reflecting on Teamwork and Class norms

For the class, we have put together some documents in reference guides about various things about teamwork. As you've done a couple of activities with your team so far, please read these and reflect on your experience and then submit what you think our class norms should be around behavior.

Do this:

Read:

1. Team Roles
2. Constructive v. Destructive Team behaviors
3. Team Meeting Types

Reflection Questions:

1. Which role do you think you'll be the best at which role do you think is growth area for you?
2. Which constructive group behavior resonates with you the most you think comes naturally to you. Which destructive behavior do you think you slip into naturally and are trying to work on?

Class norms - Step 1

in this class, we want to abide by a certain norms of behavior to maintain a positive culture and working environment. Given your experience and some of the things you've read about behaviors, please submit 3-5 norms you think we should abide by as a group. some examples of norms are:

[give examples]

You may consult any source you like, including ChatGPT (or any other other AI tools) to develop your norms.

Due:

- Submit your norms using [this form] by **midnight Thursday, Aug. 28.** (reason: we're going to collect them all and give them back to you as data).

Class norms - Step 2

By the end of the day, Friday, we will notify you when all homework is submitted and give you a reference to a spreadsheet that has everyone submitted norms. Will use that as a data set and do the following. copy and paste all submissions or upload a copy of the spreadsheet or whatever method you want to an AI tool of your choice and synthesize it down into a top 10 norms list ordered in terms of priority. It's up to you to determine the priority you can go with the consensus view you can, go based on your own personal feelings, but use the AI tool to explore analyze and question the data.

Part 2: Spreadsheet formulas and charts

FaceTime what we learned in class plus any other resources that you can find including ChatGPT or online tutorials please complete the following data tasks.

Please note there are twin objectives here. One is to give you a scenario that is common for messy data work – namely, making decision about how to create categorical data, laying the data with those categories, and then translated those categories into numbers to do further computations on them. This is a cycle of applying, and we apply some of the concepts that we learned in class. To do these tasks you may do it using whatever method you can find but can be achieved using only (a subset) of the few formulas we found in class or disgusting class average, some, men, max, if, count if.

Scenario: after your team meeting, your manager wants you to take the data a step further and categorize, crowded housing into low, medium, and high problematic areas.

Do the following:

1. Add a column to your community data spreadsheet to categorize each community area by as low medium or high problematic area for crowded housing based on some threshold you set. (hint: use if…perhaps nested!)
2. Make a bar graph that shows the count of the Number of communities that are low, medium and high. (to do this, use countif in a separate area in the spreadsheet).
3. Now add another column that converts, low, medium and high into numerical values 1,2,3 (use if again)

4. Compute the average of the numerical low/med/high values. What does this tell you? Do you think it's an "accurate" reflection of the crowded housing data? What is interesting or problematic about it? Write this down to share with your manager.

Part 3: Correlations (continue the 7 lens data scan)

The parts of the seven lens data scan that we missed last time we're about looking for correlations.

Do this:

- Pick a field that you think, correlates with crowded housing and make a scatterplot showing the relationship between the two.
- You can, of course, create multiple scatterplots relatively quickly and eyeball the correlation. Pick one that you think is relevant.
- Copy and paste the chart into the document below. Do your best to label the chart and the axis appropriately. Below the chart, write a little blurb about what the chair is showing, why you think the values might be correlated.

Part 3.5:

- For the field, you chose in part 3 pick a threshold to categorize it as low, medium and high or some other three-part categorization that makes sense.
- Produce a bar chart for this categorization as well similar to what you did in part two.

Part 4: Prioritizing

Context

A senior policymaker in your organization is preparing a citywide presentation to the mayor and council members. She wants to understand **which neighborhoods might be in the greatest need of housing intervention** She's already on board with using **crowded housing** as one of the key indicators — but she wants to **combine it with one other factor** to get a clearer picture of urgency.

"I need a simple way to show where housing stress is worst. But keep it simple. I want a grid, or maybe a map, that shows how areas fall into categories. Make it something I can explain in 30 seconds. Don't give me scatterplots with regression lines — I need to *say something actionable*."

Do This:

Use spreadsheet tools (like **sorting and filtering**) to identify:

- Neighborhoods that fall into **“High” for crowded housing**
- And that also score **“High”, “Medium”, or “Low” (whichever is most appropriate) on your chosen indicator**

Then:

- **Create a prioritized list** of neighborhoods you think the city should prioritize for attention.
- **Justify your list** in a few sentences: Why these? What combination of factors makes them stand out?
- Be ready to **show your spreadsheet view** (filtered/sorted) as evidence. (or create a separate sheet, or visualization to give your “30 second” evidence.

Optional: Advanced / Extra Credit - Make a metric

If you’re feeling confident or curious, you can go one step further:

- Convert your other indicator’s categorical labels into numbers (e.g., Low = 1, Med = 2, High = 3)
- Use a simple formula (like multiplying or averaging) to **create a metric**
- Use that metric to prioritize neighborhoods to intervene and and describe what the metric means.

- Reflect: Does this change how you’d prioritize? What does the metric capture that sorting alone doesn’t?