**Project Title: Bar Exam vs behind bars: How crime and schools play a factor on housing costs**

1. [Project Description/Outline](https://docs.google.com/document/d/1LoP-4n8t47MM8Alf8_9MgfR9wXOitiEc4cgWyH3aJq8/edit#heading=h.bttp6bv7pa86)
2. [Research Questions to Answer](https://docs.google.com/document/d/1LoP-4n8t47MM8Alf8_9MgfR9wXOitiEc4cgWyH3aJq8/edit#heading=h.gp4cyzhcuzuy)
3. [Datasets to Be Used](https://docs.google.com/document/d/1LoP-4n8t47MM8Alf8_9MgfR9wXOitiEc4cgWyH3aJq8/edit#heading=h.m91wyn7fnhnc)
4. [Description of the scope of research](https://docs.google.com/document/d/1LoP-4n8t47MM8Alf8_9MgfR9wXOitiEc4cgWyH3aJq8/edit#heading=h.wzxzs7pxb55q)
5. Methodology
6. Results

[Resources:](https://docs.google.com/document/d/1LoP-4n8t47MM8Alf8_9MgfR9wXOitiEc4cgWyH3aJq8/edit#heading=h.welduqu8mb0z)

[Project ideation](https://docs.google.com/document/d/1LoP-4n8t47MM8Alf8_9MgfR9wXOitiEc4cgWyH3aJq8/edit#heading=h.521l6lu4pe66)

[Data fetching/API integration](https://docs.google.com/document/d/1LoP-4n8t47MM8Alf8_9MgfR9wXOitiEc4cgWyH3aJq8/edit#heading=h.mh27tg23ot07)

[Data analysis: Analysis and Conclusion (20 points)](https://docs.google.com/document/d/1LoP-4n8t47MM8Alf8_9MgfR9wXOitiEc4cgWyH3aJq8/edit#heading=h.g7jclubyegem)

[Testing](https://docs.google.com/document/d/1LoP-4n8t47MM8Alf8_9MgfR9wXOitiEc4cgWyH3aJq8/edit#heading=h.k2shru38ukqa)

[Creating documentation](https://docs.google.com/document/d/1LoP-4n8t47MM8Alf8_9MgfR9wXOitiEc4cgWyH3aJq8/edit#heading=h.otpq4d7gt6vk)

[Creating the presentationSlide Deck (20 points)](https://docs.google.com/document/d/1LoP-4n8t47MM8Alf8_9MgfR9wXOitiEc4cgWyH3aJq8/edit#heading=h.4zg1kcemm2n5)

[Criteria](https://docs.google.com/document/d/1LoP-4n8t47MM8Alf8_9MgfR9wXOitiEc4cgWyH3aJq8/edit#heading=h.s26z04q4k70x)

[Project Title: The Best Place to Live based on your top Factors](https://docs.google.com/document/d/1LoP-4n8t47MM8Alf8_9MgfR9wXOitiEc4cgWyH3aJq8/edit#heading=h.sj0k2vapepqc)

**Project Description/Outline**

Buying a home in Atlanta has become more cumbersome over the last couple of years. In addition to changes in legislation and technology, much of the housing climate is ~~drowning~~ within a sea of data that is hard to process for your average home buyer. While there are many factors that are more individualized in finding a potential buyer’s “prefect choice”, there are a few factors that many will agree are almost universal. **Two of those factors are crime rates and the quality of local schools.** To help the potential home buyers navigate the housing market in Atlanta, we created a program to help process the local data on crime and local high schools' ratings to see how they may play a role in housing price decisions.

**Research Questions to Answer**

1. What is the correlation between crime rates and housing prices?
2. **What is the correlation between high school ratings and housing prices?**
3. Do these factors have any connection to each other?
   1. **Will you see areas in Atlanta where both crime rates and high school ratings are high or low?**
   2. **Are there any interesting areas found from the data?**
4. Is there any areas of recommendation for the home buyer to be aware of based on the data?
   1. Both “Good” and “Bad” territories.

**The 5-Ws**

* Who: Atlanta home buyers
* **What: Effects of housing prices based on crime and high school ratings**
* Where: Atlanta
* When: The most current data
* Why: Helping future homeowners find homes

* What are the key criteria you are searching in for?
* What is the ratings?
* What factors do we want to cross reference?
  + Commute Times Are Key
  + Local HS Schools - Kelly
  + Local Community Life
  + Neighborhood Safety
    - What type of crime are we looking for? - Ali
    - Violent Crime in a zip code
  + Nearby Parks, Walking Paths, or Dog Parks
  + Types of Restaurants
  + Financial ability - Tom
    - Ignore condos
    - Single family homes
    - Min and Max (150 - 400)
* **What type of school?**
  + **public/district HS charter or private**
  + **How many schools?  10**
  + [**https://www.atlantapublicschools.us/page/832**](https://www.atlantapublicschools.us/page/832)
  + **Do we define a radius of a certain school to look at**
* How can we cross reference school and crime?
* What area and what can we cross reference for patterns?
* What do we want to show?
* What is our outcome?
* How can we overlap schools and crimes in an area?
* How do we define the areas around the schools? How do we relate the crimes to the school area?
* What if we have two schools in the same zipcodes?
* How do we find enough data from

All zip codes, show average crime, high level zip codes

**Datasets to Be Used**

* [ZipCode DataBase](https://app.zipcodebase.com/)
  + Zipcode, city, state and distance csv posted to Github
* <https://developer.mapquest.com/documentation/geocoding-api/>
* [Great Schools](https://www.greatschools.org/gk/api-developer-resources/)
  + ​​[API documentation](https://docs.google.com/document/d/1pSe1AeZXGL01m5uG3wwRr9k4pI2Qw52xzPi2NmNrhyI/edit?usp=sharing)
* [Atlanta Crime](https://www.kaggle.com/datasets/pengchencharles/atlanta-crime-data2020?resource=download-directory)
* <https://rapidapi.com/crimeometer/api/crimeometer/>
* <https://rapidapi.com/jgentes/api/crime-data/>

**Rough Breakdown of Tasks**

* Decide on a clear Focus
* Datasets
  + Get the first dataframe of zipcodes and lat/long.
  + Save CSV of  info of  the Atlanta Area (100) in order to have a base to work with.
  + <https://www.atlantaga.gov/home/showdocument?id=3188>
* Steps to the Project
  + Each person evaluates 2 single factors against zip codes.
    1. crime,
    2. income,
    3. schools,
    4. population density,
    5. zip codes,
    6. median age,
    7. restaurants,
    8. churches,
    9. how many kids
    - Each person is responsible for 2 datasets to compare in the area codes.
  + Each person produces 2 df’s with top 10 zip codes that have that specific factor.
  + Each Person creates and writes a csv to a folder.
  + Each person puts initi[al data graphs on a slide deck](https://docs.google.com/presentation/d/1Xn_CzZBLMiWZ5ZzzO-bC0p3BPIaqww8Ik4_GQGBrVZI/edit?usp=sharing)
* Merge data
  + We merge all csv’s against the zip code, and isolate the zip code that is MAX.
  + Add weights for the most important factor?
* Graphs
  + Show the 8 factors displaying the factors

Primary Research Information

**Brainstorm Possible Project Ideas**

* Crime in/near/around schools

**Begin Data Research**

WHO- (What’s Up?)

* Questions that are interesting and what motivated us?
* Where and how did you find the data to answer the questions?

**WHAT (So What?)**

* The data exploration and cleanup process
* Analysis process

WHEN (What is the relevant time frame?)

WHERE (where is the project relevant?)

* The Atlanta metro area

WHY (Why does the audience care about the project?)

**Description of the scope of research**

**Resources:**

* City of Atlanta zip codes (not metro)  <https://www.atlantaga.gov/home/showdocument?id=3188>
* <https://developers.arcgis.com/python/samples/safe-streets-to-schools/>
* <https://www.edweek.org/leadership/school-crime-and-safety-what-a-decade-of-federal-data-show/2022/07>
* <https://www2.ed.gov/offices/OUS/PES/studies-school-violence/school-crime-pattern.pdf>
* <https://nces.ed.gov/pubs2022/2022092.pdf>
* <https://www.city-data.com/>
* <https://financesonline.com/student-crime-statistics/>
* <https://ope.ed.gov/campussafety/#/>
* <https://nces.ed.gov/fastfacts/display.asp?id=49>
* <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6545988/> crime and neighborhood of schools
* <https://opendata.atlantapd.org/>
* <https://www.atlantapd.org/i-want-to/crime-data-downloads>
* <https://www.researchgate.net/publication/258127513_Schools_Neighborhood_Risk_Factors_and_Crime>
* <https://financesonline.com/student-crime-statistics/>
* <https://sharefulton.fultoncountyga.gov/Public-Safety/Annual-Crime-Incidents-in-Atlanta-by-Type/h4ye-ggte>
* <https://www.intownelite.com/atlanta-zip-codes/> - Atlanta has **nearly 100** ZIP codes, and we asked readers last week which is their favorite.
* <https://towardsdatascience.com/student-crime-and-disruption-related-incident-prediction-in-the-us-public-schools-64d81ed129b1>
* <https://backgroundchecks.org/top-school-districts-in-georgia.html>
* <https://gbpi.org/georgia-revenue-primer-for-state-fiscal-year-2023/>
* <https://opendata.atlantapd.org/>
* <https://gist.github.com/patpohler/36c731113fd113418c0806f62cbb9e30> - Big List of Real Estate APIs
* <https://www.kaggle.com/datasets/paultimothymooney/zillow-house-price-data>

**Project ideation**

**Data fetching/API integration**

Visualizations (20 points)

* 6–8 visualizations of data (at least two per question) (10 points)
* Clear and accurate labeling of images (5 points)
* Visualizations supported with ample and precise explanation (5 points)

**Data analysis: Analysis and Conclusion (20 points)**

* Write-up summarizes major findings and implications at a professional level (5 points)
* Each question in the project proposal is answered with precise descriptions and findings (5 points)
* Findings are strongly supported with numbers and visualizations (5 points)
* Each question response is supported with a well-discerned statistical analysis from lessons (e.g., aggregation, correlation, comparison, summary statistics, sentiment analysis, and time series analysis) (5 points)

**Testing**

**Creating documentation**

Completed Analysis Uploaded to GitHub (20 points)

* Final data analysis contains ample and complete information in README file (10 points)
* Final repository is acceptable for professional quality presentation (10 points)

**Creating the presentationSlide Deck (20 points)**

* Slides are visually clean and professional (5 points)
* Slides are relevant to material (5 points)
* Slides effectively demonstrate the project (5 points)
* Slides are clear and maintain audience interest (5 points)

Group Presentation (20 points)

* All group members spoke during the presentation (5 points)
* Group was well prepared (5 points)
* Presentation is relevant to material (5 points)
* Presentation maintains audience interest (5 points)

**Criteria**

There are several criteria that can be used to evaluate school districts, including:

**Student achievement:** This can be measured by standardized test scores, graduation rates, and college attendance rates.

**Teacher quality:** This can be measured by the number of highly qualified teachers, teacher turnover rates, and teacher-to-student ratios.

**Resources:** This can include factors such as funding per student, class size, and access to technology and extracurricular activities.

**Safety and discipline:** This can include measures such as crime and violence rates, bullying, and truancy.

**Parent and community involvement:** This can include measures of parental participation in school activities and the level of community support for the schools.

**Parent/Student Surveys on Overall Experience:**

**College readiness:** This can include measures such as the number of AP classes offered, and the number of students who passed AP exams.

**Diversity and inclusion:** This can include measures such as the percentage of students from diverse backgrounds and the policies and programs in place to support them.

**Special needs support:** This can include measures such as the number of students with special needs and the resources provided to support them.

**Quality of Life:**

Ultimately, the best school district will depend on your individual priorities and what you consider most important for your child's education.

Zipcode dataframe- cvs - atlanta– ability to change - <https://www.zipcodeapi.com/> & <https://zipcode.org/city/GA/ATLANTA> & [**https://www.intownelite.com/atlanta-zip-codes/**](https://www.intownelite.com/atlanta-zip-codes/)

Zillow api- [**https://www.zillow.com/research/data/**](https://www.zillow.com/research/data/)

Waze api - commute

School api- ratings

Crime api

**Analysis and Conclusion**

* What is the distribution of Atlanta High Schools?
  + As far as the physical distribution of the high schools, there is a large grouping on the East side of the city of Atlanta along longitude -84.35.
  + The three zip codes with the most schools are 30306, 30314, and 30315.
* What is the breakdown of Atlanta High Schools?
  + Of the high schools in the Atlanta area there is a wide variety of different average scores that are passing.
  + Surprisingly, there is a large number of schools that are passing with A’s as an average.
    - It was assumed that the average letter grade would be closer to B’s
    - It should be understood that the high schools in the data set is made up of public, private, and charter instead of only one type.
* What is the correlation between high school ratings and neighborhood crime?
  + Generally speaking There is no strong correlation between the # highschools and crime rate
  + But it seems In areas where there are higher rates of crime, there are fewer high schools located.
  + It would be premature to assume that the lack of high schools has any correlation towards crime since there are other factors such as population density.
* What is the distribution of crime in Atlanta?
  + Most of the crime is centered in the down town/midtown area and toward west atlanta with a large concentration in the zip code 30318.
  + If we are referring to crime relative to population size the zip codes of 30308, 30326, 30303, and 30324 have higher rates of crime then most of Atlanta.
* What is the correlation between crime rates and housing prices?
  + Similar to schools rating, crime does seem to play a factor, and this is on clear display with the comparison between zip code: 30318 being the highest crime rate and zip code: 30327. But it seems crime plays a higher role on housing price in compare to school rating but it is not significant.
  + Most of the zip code with a decent amount of crime (+1000 in 2022) seen with similar levels of average housing cost.
    - The exception is in the north around latitude: 33.80, longitude: -84.35
  + These exceptions raise the question of how much of a pull these other factors in the Northern part of Atlanta have that they are starting to outweigh factors like crime.
* What is the distribution of average housing prices in Atlanta?
  + The average housing price seems to be centralized in the Northern part of Atlanta.
  + This was interesting since it was first assumed that a large part of the average house price distribution relative to Atlanta would contain higher prices in the center of the city since land would be more at a premium.
* Is there a trend regarding price and geographic location relative to Atlanta?
  + Most of the less wealthy areas of the city are in the Southwestern part of the city.
  + Most of the wealth is centered in the Northeastern part of the city with a huge increase around latitude: 33.85, longitude: -84.45