

Ela Gulsen

Generating Light Pollution Heat Maps of Pittsburgh

Mellon College of Science, Carnegie Mellon University

Project Information

For my research project, I've been working with Diane Turnshek and Steve Quick at CMU to generate heat maps of the light pollution in Pittsburgh. I began working on this project in early July and finished in early August.

Internship Role

Over the course of this summer, I've been working to create data visualizations of the Pittsburgh area using SQM readings taken by local residents. Although there are existing light pollution maps, most of them are only able to distinguish between general areas, such as cities, suburbs, and rural areas, with relatively low detail. By using measurements taken from citizen science projects, where anyone can upload current light pollution levels in their area from just their phone, highly accurate and detailed heatmaps can be constructed. Over the last month, I've worked on a program that would use the Seaborn Python library to create these maps.

For much of this project, I was working towards one end goal - creating a visually appealing and informative heat map of Pittsburgh. Thus, most of the following "tasks" are not

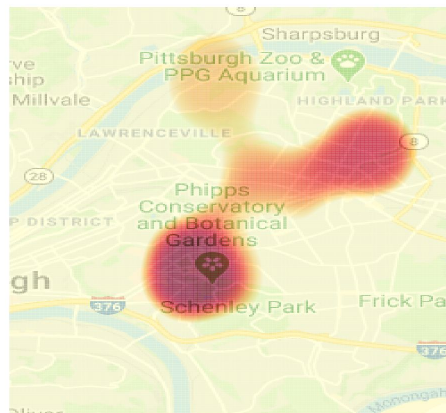
Generating Light Pollution Heat Maps of Pittsburgh

actual specific goals to accomplish but rather iterations of my work. For every version, I communicated with Professors Turnshek and Quick and we were able to find areas to improve, leading to the next version.

The project, as well as technical details about its implementation and how to run it, can be found on Github here: <https://github.com/elagulsen/skyglow-heatmaps>

Tasks

My first task involved becoming familiar with the Seaborn library and creating an initial, rough heat map with a working coordinate system and basic colors. I used the Globe at Night data sets¹, which were conveniently formatted for my benefit; however, since most submissions do not include SQM readings but are instead subjective ratings of sky brightness, only about 50 points were included in my first heat map, and nearly all of them were centered around CMU.

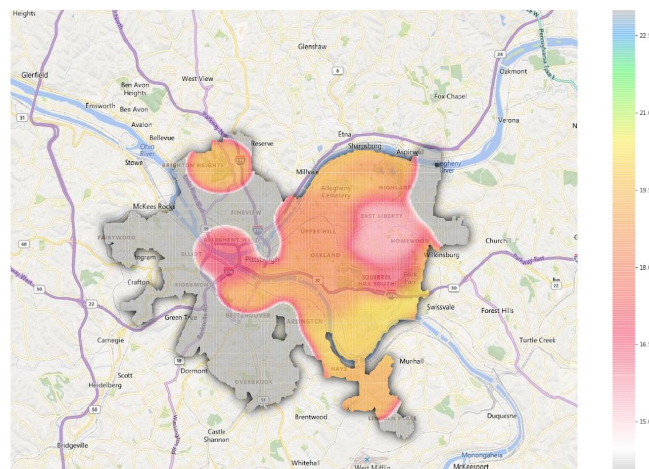


¹ <https://www.globeatnight.org/>

Generating Light Pollution Heat Maps of Pittsburgh

After generating this initial test heat map, I realized that I needed to further research topics such as light pollution standards, the accuracy of SQMs, data collection methods, etc. I also spent a significant amount of time researching alternate heat maps, including 3D ones - however, this did not end up working out.

I expanded the initial heat map to span the entirety of Pittsburgh, including a border of city limits. I also switched to the standard light pollution heat map scale, which was far more vibrant and allowed for more detail. Furthermore, I was able to find and convert Dark Sky Meter data sets ²to the appropriate format and incorporate them into the map, which brought the amount of data points included in the map to around 250.

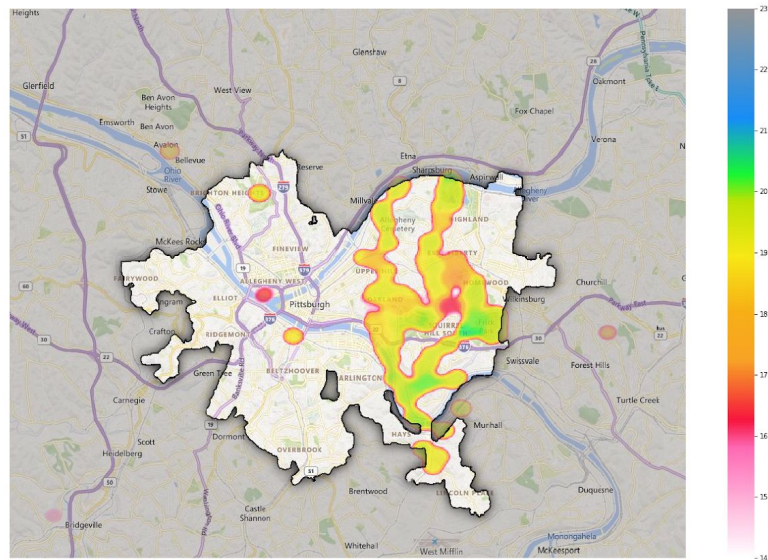


Although this map was a step in the right direction, many of the colors blended together into orange and red, which doesn't communicate much to the viewer. In order to fix this, I

² <https://www.darksky.com/>

Generating Light Pollution Heat Maps of Pittsburgh

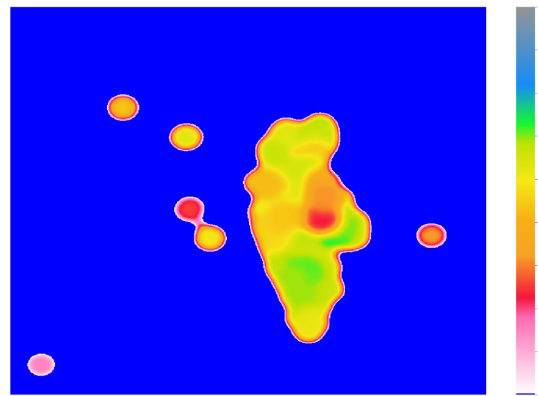
squeezed the color scale to visually distinguish highly polluted areas from lesser polluted areas. I also fixed problems such as the white borders around the colors and the dark area inside Pittsburgh.



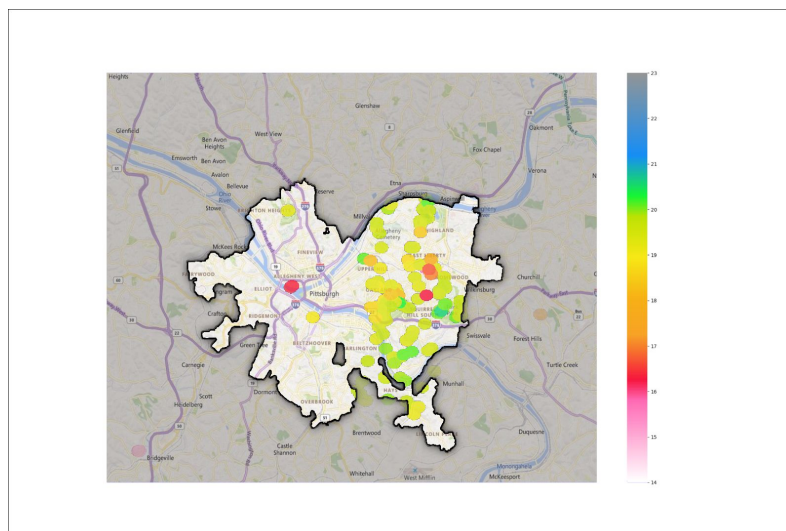
Although these changes may seem like easy cosmetic fixes, they actually required making big changes to the design of the code. Simply layering a transparent version of the heat map on top of the map of Pittsburgh (like I did before) was not working, because it produced a washed out gray background and white border (this was due to the way the blurring math was implemented). To fix this, I added a new step to the process: I first generated a heat map that had a bright blue background, and then wrote code to go through the entire image to remove all

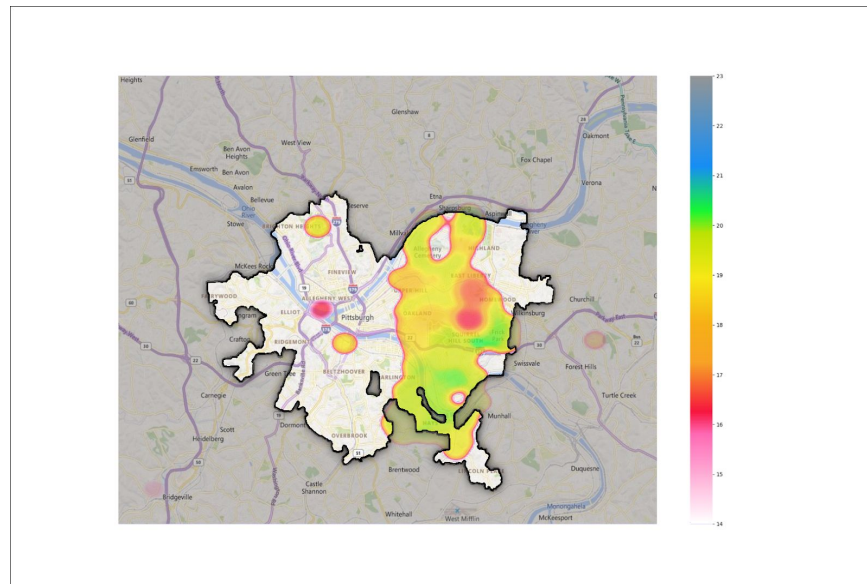
Generating Light Pollution Heat Maps of Pittsburgh

traces of blue and white from it. I then layered this image, with a transparent background, over the map of Pittsburgh to produce a nice, vibrant final product.



The final changes I made revolved around optimizing code and documenting it, rather than producing new images. I switched from Jupyter Notebook (which I had been testing on) to regular Python, and uploaded my code to Github ([link above](#)). I focused on making my code faster, explaining my processes so that others might be able to revisit it in the future, and making it more user-friendly.



Generating Light Pollution Heat Maps of Pittsburgh

I was not able to accomplish everything that I initially hoped I could, due to a couple of issues. One of these was the lack of data available; there simply aren't enough readings to fill up a detailed map of Pittsburgh, and there aren't many people willing to go out and take measurements. In the future, it would be great to take readings of downtown and western Pittsburgh, of which readings are noticeably absent. In particular, if many readings could be taken over a long period of time, multiple heat maps could be made to show the changes in light pollution levels over time.

Another problem I encountered was that I was limited by my experience in data science. Before the project, I had little experience working with data science in Python, and had to learn how to use many of the libraries involved in this project. Although I briefly experimented with different types of figures, like 3D graphs, I ultimately decided that it would be better to focus my

time on optimizing the 2D maps I had already created. However, creating alternate data visualizations would be a great consideration if this project were to be continued in the future.

Experience Description

When the research project was introduced, I knew that it would be the perfect opportunity to try working on a practical project that incorporated math/computer science. When I looked through the project list, however, I saw that many of the projects seemed very large-scale and complicated, and involved skills that I was totally unfamiliar with, like computer vision. When I talked to Professors Turnshek and Quick, though, they offered me the perfect opportunity. Generating heat maps of Pittsburgh was a relatively small-scale and specific task, yet interesting and non-trivial.

When I got started on the project, I was in complete free-reign of the programming and implementation, which was overwhelming. A classmate, Prabh, pointed me in the right direction and sent me resources that dealt with the libraries I would be using to generate the core of the heatmap. With those in mind, I got started on the programming, which was entirely self-directed.

After every new iteration of design I produced, I talked to Diane and Steve a lot about the visuals of the images I was generating. Through countless back and forth emails, they provided the feedback that let me improve my work. In particular, Diane provided insight to the astronomical perspective that I didn't have. She introduced me to resources like the Dark Sky

Meter website, the standardized light pollution heatmap color scheme, and through her I even attended a Pittsburgh IDA meeting (which was fun!).

Overall, I had a great experience. Since I'm relatively inexperienced with practical projects like this, it was great to be able to take my time and really focus on learning and experimenting. I liked the fact that I was working towards one end goal, and every iteration of my work was singularly focused on getting to that point. I'm not sure if I would have gotten the same out of a project where I had to work on something more large-scale.

Knowledge Gained

Over the course of this project, I've become comfortable using several Python libraries used in data science, like Seaborn, Matplotlib, Pandas, and Numpy. I've also picked up a lot of astronomy knowledge, through researching light pollution levels and reading papers. But more importantly than the technical skills I learned, I've gained a lot of experience in directing my own work and designing my own programs.

Much of the experience gained is due to the independent nature of this project. Although I've taken a few computer science classes before, the skills taught by those courses were drastically different from the ones I developed over the course of this project. Assignments in classes involved solving already-defined problems, with clear expectations, straightforward

Generating Light Pollution Heat Maps of Pittsburgh

designs, and, often, with tests provided. Yet this project necessitated that I direct my own learning and set my own deadlines. When I had questions or problems, I had to do my own research instead of relying on others to tell me what to do, and I had to assess the quality of my own work instead of submitting it to be graded. This project has allowed me to grow not only as a programmer, but also as a project manager and independent thinker.