

NYC Opportunity EDS – Take-home Assessment for Data Curator/Scientist

The purpose of this take-home exam is to give you an opportunity to explore NYC-related datasets, work through a real issue the City has been facing.

Use Case - Effective eviction prevention/outreach efforts

Let's say that you are a Data Curator/Scientist at NYC Opportunity, and the [Mayor's Office to Protect Tenants](#) (created by [Executive Order No. 39](#)) asks you to analyze to help improve outreach efforts to those who may be at risk of eviction.

Task 1: Director Stats

The Executive Director of the Mayor's Office to Protect Tenants is interested in providing stats on Housing Court activity in New York City for a presentation for a Deputy Mayor.

The Housing Data Coalition makes housing-related datasets easily available to the public. See their GitHub page [here](#) (instructions to access the dataset can be found here). Using the OCA Housing Court Records [dataset](#), answer the following summary statistics.

1. How have many total cases for residential housing there been in each NYC borough?
2. Which month of the year sees the highest number of housing court cases?
3. What are the top 10 NYC ZIP codes with residential court cases before March 2020?
4. Are there any neighborhoods (based on ZIP codes) that show relatively more cases after March 2020 compared to pre-pandemic?

Task 2: Tenant Outreach

Develop a set of recommendations for the Mayor's Office to Protect Tenants to improve outreach efforts to tenants who might be at risk of appearing in court or eviction. You can start with the OCA dataset you use for Task 1, [eviction data](#) from NYC Open Data, [NYC PLUTO data](#), census data, or other data sets. You are not required to use all suggested data sets, but also we encourage you to use any methodologies or additional datasets that would help you answer this question.

We understand that the analysis request can be a months-long project, and you only have a few hours. While you can't explore everything and answer all of the questions in your mind, please remember that we're more interested in seeing your thinking process, and there's no "right" answer.

Task 3: Data Documentation Exercise

These are open-ended questions and intended to help understand your approach to data documentation.

1. What do you think are ideal sources for data documentation?
2. How can data documentation affect your analysis of the data?
3. Why is it important to have thorough data documentation?

Please include the following documents in your response:

1. A one-page of answers from Task 1: Director Stats using the Housing court data
2. A one paragraph memo of key findings and recommendations from your analysis in Task 2: Tenant Outreach. You can include a visualization to support your key findings.
3. A one-page (maximum) technical note to show your analysis from Task 2: Tenant Outreach. It can follow the outline below:
 - a. Define the problem: How do you turn their questions into your own problem statement?
 - b. Analysis Process: What methodology, models, analysis, and/or techniques you use for the analysis? Why did you choose them?
 - c. Caveat/Next Steps: Please feel free to add the following steps if you have more time and/or more data!
4. The code you're using for the analysis. We accept all text format (.r or .py) or in a jupyter notebook (.ipynb)
5. A one-page response to the Task 3: Data Documentation Exercise

Please submit all your documents to analytics@nycopportunity.nyc.gov 48 hours after receipt of this email.