Final Project - Proposal

The data sets I chose are related to employee salaries of Baltimore city for the years 2016-2019.

The website URL: <https://data.world/datasets/employees>

The URLs to access the datasets:

2019: <https://query.data.world/s/wrxyalft7tm2qlg7soxjv75tw5a26o>

2018: <https://query.data.world/s/kxh2l22bf3id3uj2uutljvwzcbfhv6>

2017: <https://query.data.world/s/mfg4mszxwdgs4qw7m2kjmqq5wpdw7x>

2016: <https://query.data.world/s/bbh4zuwcfl6ugi2iahvh5m74sqrbik>

* A week ago, I was surfing about the salaries of employees here in USA. And now, I think it’ll be interesting to find insights like which profession is highly paid and to observe the change in their salaries over years. So, I searched for “Baltimore employee salaries” and found a bunch of links. “data.world” has multiple files with yearly salary data, so I chose this website.
* Data size:
  + 2019: 7 columns\*13812 rows – Size 1478 Kb
  + 2018: 7 columns\*13684 rows – Size 1540 Kb
  + 2017: 7 columns\*13484 rows – Size 1517 Kb
  + 2016: 7 columns\*13819 rows – Size 1556 Kb
* The data is free to download, and I did not violate any terms of service.
* The data has employee details like name, designation, hired date and salary. After combining the 4 datasets I’ll be able to analyze data of four consecutive years containing salaries of employees. By doing this I can closely observe the small changes in professions. Also, I can see the consistency of employee count in each field.

->Hypothesis from the data

* The number of salary hikes will be more than deductions, when 2016 to 2019 datasets are compared.
* Employees in Police department might have the highest gross income.

-> Questions

* Which profession is highest paid?
* Is the same profession paid highest in all the four years?
* Which profession has a drastic reduce in remuneration?