Tenth Meeting Presentation

Daniel L.

4/14/2022

The American Community Survey is provided to residents every year in the United States. For data regarding the population of NYC, the NYC Department of City Planning has provided population data for the years of 2005 through 2019. Originally, I was going to attempt to query the data directly from the ACS website. However, the data provided by the NYC agency is easier and better than cleaning the data. I have decided to investigate certain characteristics from the "Demographic" and "Economic" tables for the years of 2017 through 2019. For the years of 2017, and 2018, they are 3 year estimates. As mentioned, this data is not in a monthly timescale, but rather yearly estimates. What makes me hesitant to use these as forecasting variables is that these are estimates/counts. Which will most likely lead to a count regression approach that I would not like to take.

2019 ACS data

With regards to the Demographic data for 2019, the fields that I believe were worth looking at were total estimates, by borough, of the "Total Population", "Hispanic/Latino" population, "Black and African alone" population, and "Asian Alone" population.

```
demo_2019_acs5yr_cdta %>%
  group_by(Borough) %>%
  select(Borough,
           ends_with("E",
                     ignore.case = FALSE)) %>%
  summarise(across(where(is.numeric), total, .names = "{.col}_{.fn}")) %>%
  select(Borough,c(Pop_1E_total, Hsp2E_total, BlNHE_total, Asn1RcE_total))
```

```
## # A tibble: 5 x 5
##   Borough       Pop_1E_total Hsp2E_total BlNHE_total Asn1RcE_total
##   <chr>                <dbl>       <dbl>       <dbl>         <dbl>
## 1 Bronx              1438758      809507      417737         52883
## 2 Brooklyn           2589967      493231      776274        307082
## 3 Manhattan          1621398      413634      201694        198758
## 4 Queens             2294241      643903      397341        583642
## 5 Staten Island       474277       87443       44420         44225
```

For the 2019 economic data, the fields that I believe were worth looking at were the total estimates, by borough, of the "Population 16 years and over who are employed", ""Population 16 years and over who are unemployed", "Workers 16 years and over who use public transportation", "Workers 16 years and over who work from home", and "total households".

```
econ_2019_acs5yr_cdta %>%
  group_by(Borough) %>%
  select(Borough,
         ends_with("E",
                   ignore.case = FALSE)) %>%
  summarise(across(where(is.numeric), total, .names = "{.col}_{.fn}")) %>%
  select(Borough,c(CvEm16pl1E_total, CvLFUEm1E_total, CW_PbTrnsE_total,
                   CW_WrkdHmE_total, HH2E_total))
```

```
## # A tibble: 5 x 6
##   Borough     CvEm16pl1E_total CvLFUEm1E_total CW_PbTrnsE_total CW_WrkdHmE_total
##   <chr>                  <dbl>           <dbl>            <dbl>            <dbl>
## 1 Bronx                 605864           67333           350854            17674
## 2 Brooklyn             1227022           80619           733656            54724
## 3 Manhattan             900945           49284           518789            63678
## 4 Queens               1134877           67144           563503            29978
## 5 Staten Isl~           214342           10287            62621             5638
## # ... with 1 more variable: HH2E_total <dbl>
```

2018 ACS data

```
demo_2018_acs5yr_puma %>%
  group_by(Borough) %>%
  select(Borough,
           ends_with("E",
                     ignore.case = FALSE)) %>%
  summarise(across(where(is.numeric), total, .names = "{.col}_{.fn}")) %>%
  select(Borough,c(Pop_1E_total, Hsp2E_total, BlNHE_total, Asn1RcE_total))
```

```
## # A tibble: 5 x 5
##   Borough       Pop_1E_total Hsp2E_total BlNHE_total Asn1RcE_total
##   <chr>                <dbl>       <dbl>       <dbl>         <dbl>
## 1 Bronx              1437872      803636      421275         52388
## 2 Brooklyn           2600747      499279      787705        306829
## 3 Manhattan          1632480      423683      203849        196347
## 4 Queens             2298611      643563      396392        580725
## 5 Staten Island       474101       86976       43906         41427
```

```
econ_2018_acs5yr_puma %>%
  group_by(Borough) %>%
  select(Borough,
         ends_with("E",
                   ignore.case = FALSE)) %>%
  summarise(across(where(is.numeric), total, .names = "{.col}_{.fn}")) %>%
  select(Borough,c(CvEm16pl1E_total, CvLFUEm1E_total, CW_PbTrnsE_total,
                   CW_WrkdHmE_total, HH2E_total))
```

```
## # A tibble: 5 x 6
##   Borough     CvEm16pl1E_total CvLFUEm1E_total CW_PbTrnsE_total CW_WrkdHmE_total
##   <chr>                  <dbl>           <dbl>            <dbl>            <dbl>
```

```
## 1 Bronx                595657          70250          346240           18258
## 2 Brooklyn            1217624          91531          729737           52118
## 3 Manhattan            897040          54298          521777           60494
## 4 Queens              1131583          74393          563428           28423
## 5 Staten Isl~          211293          10872           61905            5301
## # ... with 1 more variable: HH2E_total <dbl>
```

2017 ACS data

```
demo_2017_acs5yr_puma %>%
  group_by(Borough) %>%
  select(Borough,
           ends_with("E",
                     ignore.case = FALSE)) %>%
  summarise(across(where(is.numeric), total, .names = "{.col}_{.fn}")) %>%
  select(Borough,c(Pop_1E_total, Hsp2E_total, BlNHE_total, Asn1RcE_total))
```

```
## # A tibble: 5 x 5
##   Borough       Pop_1E_total Hsp2E_total BlNHE_total Asn1RcE_total
##   <chr>                <dbl>       <dbl>       <dbl>         <dbl>
## 1 Bronx              1455846      810549      427414         53413
## 2 Brooklyn           2635121      508365      805388        313891
## 3 Manhattan          1653877      431249      206945        198496
## 4 Queens             2339328      654793      401562        592267
## 5 Staten Island       475948       86540       44600         40287
```

```
econ_2017_acs5yr_puma %>%
  group_by(Borough) %>%
  select(Borough,
         ends_with("E",
                   ignore.case = FALSE)) %>%
  summarise(across(where(is.numeric), total, .names = "{.col}_{.fn}")) %>%
  select(Borough,c(CvEm16pl1E_total, CvLFUEm1E_total, CW_PbTrnsE_total,
                   CW_WrkdHmE_total, HH2E_total))
```

```
## # A tibble: 5 x 6
##   Borough     CvEm16pl1E_total CvLFUEm1E_total CW_PbTrnsE_total CW_WrkdHmE_total
##   <chr>                  <dbl>           <dbl>            <dbl>            <dbl>
## 1 Bronx                 593858           77907           347158            18039
## 2 Brooklyn             1217976          106976           735534            50622
## 3 Manhattan             906389           59771           528413            59886
## 4 Queens               1142378           85174           573952            28667
## 5 Staten Isl~           212253           12444            62675             5323
## # ... with 1 more variable: HH2E_total <dbl>
```