Floodzone Data Analysis - County Level

Mattingly, Peter

10 September 2020

Contents

Turn Off Scientific Notation

```
options(scipen=15)
print(1e15)
```

```
## [1] 1000000000000000
```

Lood Floodzone Data

```
# Import NY Floodzone Data
ny_floodzone <- read_excel("~/Desktop/Thesis Analysis/Data/NY_FloodzoneData.xlsx",
                           sheet = "County Data")
ny_floodzone <- ny_floodzone %>%
  rename_all(. %>% tolower %>% gsub(" ", "_", .))
```

Population

```
pop_vars <- c("geo_type","geo_name","geo_id","stateusps",
              "pop_tot","pop_100year","pop_anyyear")

ny_floodzone_pop <- ny_floodzone %>%
  select(pop_vars) %>%
  gather(subject, population, pop_tot:pop_anyyear) %>%
  separate(subject, c("variable", "universe")) %>%
  select(-variable) %>%
  mutate(universe = replace(universe, universe == "tot", "Total"),
         universe = replace(universe, universe == "100year", "100-Year"),
         universe = replace(universe, universe == "anyyear", "Combined")) %>%
  filter(population != 0)
```

```
## Note: Using an external vector in selections is ambiguous.
## i Use `all_of(pop_vars)` instead of `pop_vars` to silence this message.
## i See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This message is displayed once per session.
```

```
head(ny_floodzone_pop)
```

```
## # A tibble: 6 x 6
##   geo_type geo_name geo_id stateusps universe population
##   <chr>    <chr>    <chr>  <chr>     <chr>         <dbl>
## 1 County   Albany   36001  NY        Total        307463
```

```
## 2 County   Bronx    36005  NY        Total       1428357
## 3 County   Cayuga   36011  NY        Total         79173
## 4 County   Chenango 36017  NY        Total         49549
## 5 County   Clinton  36019  NY        Total         81685
## 6 County   Cortland 36023  NY        Total         49043
```

Housing

```
housing_vars <- c("geo_type","geo_name","geo_id","stateusps",
                  "hu_100year","hu_anyyear","hu_tot")

ny_floodzone_housing <- ny_floodzone %>%
  select(housing_vars) %>%
  rename_all(funs(stringr::str_replace_all(., c("hu_"), ""))) %>%
  gather(universe, housing_total, "100year":tot) %>%
  mutate(universe = replace(universe, universe == "tot", "Total"),
         universe = replace(universe, universe == "100year", "100-Year"),
         universe = replace(universe, universe == "anyyear", "Combined")) %>%
  filter(housing_total != 0)
```

```
## Note: Using an external vector in selections is ambiguous.
## i Use `all_of(housing_vars)` instead of `housing_vars` to silence this message.
## i See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This message is displayed once per session.

## Warning: `funs()` is deprecated as of dplyr 0.8.0.
## Please use a list of either functions or lambdas:
##
##   # Simple named list:
##   list(mean = mean, median = median)
##
##   # Auto named with `tibble::lst()`:
##   tibble::lst(mean, median)
##
##   # Using lambdas
##   list(~ mean(., trim = .2), ~ median(., na.rm = TRUE))
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_warnings()` to see where this warning was generated.
```

```
head(ny_floodzone_housing)
```

```
## # A tibble: 6 x 6
##   geo_type geo_name geo_id stateusps universe housing_total
##   <chr>    <chr>    <chr>  <chr>     <chr>           <dbl>
## 1 County   Albany   36001  NY        100-Year         3603.
## 2 County   Bronx    36005  NY        100-Year         4966.
## 3 County   Cayuga   36011  NY        100-Year         2216.
## 4 County   Chenango 36017  NY        100-Year         3347.
## 5 County   Clinton  36019  NY        100-Year         2218.
## 6 County   Cortland 36023  NY        100-Year         1932.
```

Race

```
race_vars <- c("geo_type","geo_name","geo_id","stateusps",
               "shr_pop_asian_fp_100","shr_pop_asian_fp_any",
               "shr_pop_asian_county","shr_pop_black_fp_100",
               "shr_pop_black_fp_any","shr_pop_black_county",
               "shr_pop_hispanic_fp_100","shr_pop_hispanic_fp_any",
               "shr_pop_hispanic_county","shr_pop_white_fp_100",
               "shr_pop_white_fp_any","shr_pop_white_county")

ny_floodzone_race <- ny_floodzone %>%
  select(race_vars) %>%
  rename_all(funs(stringr::str_replace_all(., c("shr_pop_"), ""))) %>%
  rename_all(funs(stringr::str_replace_all(., c("_fp"), ""))) %>%
  gather(subject, race_share, asian_100:white_county) %>%
  separate(subject, c("race", "universe")) %>%
  mutate(race = str_to_title(race),
         universe = replace(universe, universe == "tract", "Total"),
         universe = replace(universe, universe == "100", "100-Year"),
         universe = replace(universe, universe == "any", "Combined")) %>%
  filter(race_share != "NA")
```

```
## Note: Using an external vector in selections is ambiguous.
## i Use `all_of(race_vars)` instead of `race_vars` to silence this message.
## i See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This message is displayed once per session.
```

```
head(ny_floodzone_race)
```

```
## # A tibble: 6 x 7
##   geo_type geo_name geo_id stateusps race  universe race_share
##   <chr>    <chr>    <chr>  <chr>     <chr> <chr>         <dbl>
## 1 County   Albany   36001  NY        Asian 100-Year    0.0563
## 2 County   Bronx    36005  NY        Asian 100-Year    0.0316
## 3 County   Cayuga   36011  NY        Asian 100-Year    0.00383
## 4 County   Chenango 36017  NY        Asian 100-Year    0.00552
## 5 County   Clinton  36019  NY        Asian 100-Year    0.0278
## 6 County   Cortland 36023  NY        Asian 100-Year    0.00872
```

Income

```
income_vars <- c("geo_type","geo_name","geo_id","stateusps",
                 "shr_hu_incless25_fp_100","shr_hu_incless25_fp_any",
                 "shr_hu_incless25_county","shr_hu_inc25to50_fp_100",
                 "shr_hu_inc25to50_fp_any","shr_hu_inc25to50_county",
                 "shr_hu_inc50to75_fp_100","shr_hu_inc50to75_fp_any",
                 "shr_hu_inc50to75_county","shr_hu_inc75up_fp_100",
                 "shr_hu_inc75up_fp_any","shr_hu_inc75up_county")

ny_floodzone_income <- ny_floodzone %>%
  select(income_vars) %>%
  rename_all(funs(stringr::str_replace_all(., c("shr_hu_"), ""))) %>%
  rename_all(funs(stringr::str_replace_all(., c("_fp"), ""))) %>%
  gather(subject, income_share, incless25_100:inc75up_county) %>%
```

```
  separate(subject, c("income", "universe")) %>%
  mutate(income = replace(income, income == "incless25", "<$25k"),
         income = replace(income, income == "inc25to50", "$25k-$49,999"),
         income = replace(income, income == "inc50to75", "$50k-$74,999"),
         income = replace(income, income == "inc75up", "$75k+"),
         universe = replace(universe, universe == "tract", "Total"),
         universe = replace(universe, universe == "100", "100-Year"),
         universe = replace(universe, universe == "any", "Combined")) %>%
  filter(income_share != "NA")
```

```
## Note: Using an external vector in selections is ambiguous.
## i Use `all_of(income_vars)` instead of `income_vars` to silence this message.
## i See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This message is displayed once per session.
```

```
head(ny_floodzone_income)
```

```
## # A tibble: 6 x 7
##   geo_type geo_name geo_id stateusps income universe income_share
##   <chr>    <chr>    <chr>  <chr>     <chr>  <chr>           <dbl>
## 1 County   Albany   36001  NY        <$25k  100-Year        0.241
## 2 County   Bronx    36005  NY        <$25k  100-Year        0.241
## 3 County   Cayuga   36011  NY        <$25k  100-Year        0.185
## 4 County   Chenango 36017  NY        <$25k  100-Year        0.277
## 5 County   Clinton  36019  NY        <$25k  100-Year        0.242
## 6 County   Cortland 36023  NY        <$25k  100-Year        0.240
```

Housing Tenure

```
tenure_vars <- c("geo_type","geo_name","geo_id","stateusps",
                 "shr_hu_renter_fp_100","shr_hu_renter_fp_any",
                 "shr_hu_renter_county","shr_hu_owner_fp_100",
                 "shr_hu_owner_fp_any","shr_hu_owner_county")

ny_floodzone_tenure <- ny_floodzone %>%
  select(tenure_vars) %>%
  rename_all(funs(stringr::str_replace_all(., c("shr_hu_"), ""))) %>%
  rename_all(funs(stringr::str_replace_all(., c("_fp"), ""))) %>%
  gather(subject, tenure_share, renter_100:owner_county) %>%
  separate(subject, c("tenure", "universe")) %>%
  mutate(tenure = str_to_title(tenure),
         universe = replace(universe, universe == "tract", "Total"),
         universe = replace(universe, universe == "100", "100-Year"),
         universe = replace(universe, universe == "any", "Combined")) %>%
  filter(tenure_share != "NA")
```

```
## Note: Using an external vector in selections is ambiguous.
## i Use `all_of(tenure_vars)` instead of `tenure_vars` to silence this message.
## i See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This message is displayed once per session.
```

```r
head(ny_floodzone_tenure)
```

```
## # A tibble: 6 x 7
##   geo_type geo_name geo_id stateusps tenure universe tenure_share
##   <chr>    <chr>    <chr>  <chr>     <chr>  <chr>           <dbl>
## 1 County   Albany   36001  NY        Renter 100-Year        0.569
## 2 County   Bronx    36005  NY        Renter 100-Year        0.481
## 3 County   Cayuga   36011  NY        Renter 100-Year        0.172
## 4 County   Chenango 36017  NY        Renter 100-Year        0.297
## 5 County   Clinton  36019  NY        Renter 100-Year        0.345
## 6 County   Cortland 36023  NY        Renter 100-Year        0.337
```

Building Type

```r
unit_vars <- c("geo_type","geo_name","geo_id","stateusps",
               "shr_hu_1fam_fp_100","shr_hu_1fam_fp_any",
               "shr_hu_1fam_county","shr_hu_2to4fam_fp_100",
               "shr_hu_2to4fam_fp_any","shr_hu_2to4fam_county",
               "shr_hu_5to19fam_fp_100","shr_hu_5to19fam_fp_any",
               "shr_hu_5to19fam_county","shr_hu_20up_fp_100",
               "shr_hu_20up_fp_any","shr_hu_20up_county",
               "shr_hu_other_fp_100","shr_hu_other_fp_any",
               "shr_hu_other_county")

ny_floodzone_unit <- ny_floodzone %>%
  select(unit_vars) %>%
  rename_all(funs(stringr::str_replace_all(., c("shr_hu_"), ""))) %>%
  rename_all(funs(stringr::str_replace_all(., c("_fp"), ""))) %>%
  gather(subject, unit_share, "1fam_100":other_county) %>%
  separate(subject, c("unit", "universe")) %>%
  mutate(unit = replace(unit, unit == "1fam", "Single-Family"),
         unit = replace(unit, unit == "2to4fam", "2-4"),
         unit = replace(unit, unit == "5to19fam", "5-19"),
         unit = replace(unit, unit == "20up", "20+"),
         unit = replace(unit, unit == "other", "Other"),
         universe = replace(universe, universe == "tract", "Total"),
         universe = replace(universe, universe == "100", "100-Year"),
         universe = replace(universe, universe == "any", "Combined")) %>%
  filter(unit_share != "NA")
```

```
## Note: Using an external vector in selections is ambiguous.
## i Use `all_of(unit_vars)` instead of `unit_vars` to silence this message.
## i See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This message is displayed once per session.
```

```r
head(ny_floodzone_unit)
```

```
## # A tibble: 6 x 7
##   geo_type geo_name geo_id stateusps unit          universe unit_share
##   <chr>    <chr>    <chr>  <chr>     <chr>         <chr>         <dbl>
## 1 County   Albany   36001  NY        Single-Family 100-Year      0.366
```

```
## 2 County   Bronx    36005  NY        Single-Family 100-Year      0.433
## 3 County   Cayuga   36011  NY        Single-Family 100-Year      0.756
## 4 County   Chenango 36017  NY        Single-Family 100-Year      0.594
## 5 County   Clinton  36019  NY        Single-Family 100-Year      0.605
## 6 County   Cortland 36023  NY        Single-Family 100-Year      0.641
```

Poverty

```
poverty_vars <- c("geo_type","geo_name","geo_id","stateusps",
                  "shr_pop_poverty_fp_100","shr_pop_poverty_fp_any",
                  "shr_pop_poverty_county")

ny_floodzone_poverty <- ny_floodzone %>%
  select(poverty_vars) %>%
  rename_all(funs(stringr::str_replace_all(., c("shr_pop_poverty_"), ""))) %>%
  rename_all(funs(stringr::str_replace_all(., c("fp_"), ""))) %>%
  gather(subject, poverty_share, "100":county) %>%
  separate(subject, c("universe")) %>%
  mutate(universe = replace(universe, universe == "tract", "Total"),
         universe = replace(universe, universe == "100", "100-Year"),
         universe = replace(universe, universe == "any", "Combined")) %>%
  filter(poverty_share != "NA")
```

```
## Note: Using an external vector in selections is ambiguous.
## i Use `all_of(poverty_vars)` instead of `poverty_vars` to silence this message.
## i See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This message is displayed once per session.
```

```
head(ny_floodzone_poverty)
```

```
## # A tibble: 6 x 6
##   geo_type geo_name geo_id stateusps universe poverty_share
##   <chr>    <chr>    <chr>  <chr>     <chr>            <dbl>
## 1 County   Albany   36001  NY        100-Year        0.143
## 2 County   Bronx    36005  NY        100-Year        0.167
## 3 County   Cayuga   36011  NY        100-Year        0.0984
## 4 County   Chenango 36017  NY        100-Year        0.171
## 5 County   Clinton  36019  NY        100-Year        0.156
## 6 County   Cortland 36023  NY        100-Year        0.177
```

Subsidized Housing

```
subsidized_vars <- c("geo_type","geo_name","geo_id","stateusps",
                  "ph_units_100year","ph_units_anyyear",
                  "sub_housing_units_100year","sub_housing_units_anyyear")

ny_floodzone_subsidized <- ny_floodzone %>%
  select(subsidized_vars) %>%
  rename_all(funs(stringr::str_replace_all(., c("housing_"), ""))) %>%
  rename_all(funs(stringr::str_replace_all(., c("units_"), ""))) %>%
  gather(subject, subsidized_units, ph_100year:sub_anyyear) %>%
  separate(subject, c("subsidized", "universe")) %>%
```

```
  mutate(subsidized = replace(subsidized, subsidized == "ph", "Public Housing"),
         subsidized = replace(subsidized, subsidized == "sub", "Subsidized"),
         universe = replace(universe, universe == "100year", "100-Year"),
         universe = replace(universe, universe == "anyyear", "Combined")) %>%
  filter(subsidized_units != "NA")
```

```
## Note: Using an external vector in selections is ambiguous.
## i Use `all_of(subsidized_vars)` instead of `subsidized_vars` to silence this message.
## i See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This message is displayed once per session.
```

```
head(ny_floodzone_subsidized)
```

```
## # A tibble: 6 x 7
##   geo_type geo_name  geo_id stateusps subsidized  universe subsidized_units
##   <chr>    <chr>     <chr>  <chr>     <chr>       <chr>               <dbl>
## 1 County   Albany    36001  NY        Public Hou~ 100-Year              382
## 2 County   Kings     36047  NY        Public Hou~ 100-Year             5578
## 3 County   Nassau    36059  NY        Public Hou~ 100-Year              845
## 4 County   New York  36061  NY        Public Hou~ 100-Year             9243
## 5 County   Queens    36081  NY        Public Hou~ 100-Year             3411
## 6 County   Westches~ 36119  NY        Public Hou~ 100-Year               96
```

Building Age

```
years_vars <- c("geo_type","geo_name","geo_id","stateusps",
                "shr_hu_builtpre60_fp_100","shr_hu_builtpre60_fp_any",
                "shr_hu_builtpre60_county","shr_hu_built6079_fp_100",
                "shr_hu_built6079_fp_any","shr_hu_built6079_county",
                "shr_hu_built8099_fp_100","shr_hu_built8099_fp_any",
                "shr_hu_built8099_county","shr_hu_built00s_fp_100",
                "shr_hu_built00s_fp_any","shr_hu_built00s_county")

ny_floodzone_age <- ny_floodzone %>%
  select(years_vars) %>%
  rename_all(funs(stringr::str_replace_all(., c("shr_hu_"), ""))) %>%
  rename_all(funs(stringr::str_replace_all(., c("fp_"), ""))) %>%
  gather(subject, age_share, builtpre60_100:built00s_county) %>%
  separate(subject, c("housing_age", "universe")) %>%
  mutate(housing_age = replace(housing_age, housing_age == "builtpre60", "60+"),
         housing_age = replace(housing_age, housing_age == "built6079", "41-60"),
         housing_age = replace(housing_age, housing_age == "built8099", "21-40"),
         housing_age = replace(housing_age, housing_age == "built00s", "<21"),
         universe = replace(universe, universe == "tract", "Total"),
         universe = replace(universe, universe == "100", "100-Year"),
         universe = replace(universe, universe == "any", "Combined")) %>%
  filter(age_share != "NA")
```

```
## Note: Using an external vector in selections is ambiguous.
## i Use `all_of(years_vars)` instead of `years_vars` to silence this message.
## i See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This message is displayed once per session.
```

```
head(ny_floodzone_age)
```

```
## # A tibble: 6 x 7
##   geo_type geo_name geo_id stateusps housing_age universe age_share
##   <chr>    <chr>    <chr>  <chr>     <chr>       <chr>        <dbl>
## 1 County   Albany   36001  NY        60+         100-Year     0.547
## 2 County   Bronx    36005  NY        60+         100-Year     0.496
## 3 County   Cayuga   36011  NY        60+         100-Year     0.462
## 4 County   Chenango 36017  NY        60+         100-Year     0.535
## 5 County   Clinton  36019  NY        60+         100-Year     0.398
## 6 County   Cortland 36023  NY        60+         100-Year     0.653
```

Buildings 60+ Years

```
age_60_vars <- c("geo_type","geo_name","geo_id","stateusps",
              "shr_hu_pre60_fp_100_1fam","shr_hu_pre60_fp_any_1fam",
              "shr_hu_pre60_fp_100_2to4fam","shr_hu_pre60_fp_any_2to4fam",
              "shr_hu_pre60_fp_100_5to19fam","shr_hu_pre60_fp_any_5to19fam",
              "shr_hu_pre60_fp_100_20up","shr_hu_pre60_fp_any_20up",
              "shr_hu_pre60_fp_100_other","shr_hu_pre60_fp_any_other")

ny_floodzone_60years <- ny_floodzone %>%
  select(age_60_vars) %>%
  rename_all(funs(stringr::str_replace_all(., c("shr_hu_pre60_"), ""))) %>%
  rename_all(funs(stringr::str_replace_all(., c("fp_"), ""))) %>%
  gather(subject, age_60_share, "100_1fam":any_other) %>%
  separate(subject, c("universe", "unit")) %>%
  mutate(universe = replace(universe, universe == "tract", "Total"),
         universe = replace(universe, universe == "100", "100-Year"),
         universe = replace(universe, universe == "any", "Combined"),
         unit = replace(unit, unit == "1fam", "Single-Family"),
         unit = replace(unit, unit == "2to4fam", "2-4"),
         unit = replace(unit, unit == "5to19fam", "5-19"),
         unit = replace(unit, unit == "20up", "20+"),
         unit = replace(unit, unit == "other", "Other")) %>%
  filter(age_60_share != "NA")
```

```
## Note: Using an external vector in selections is ambiguous.
## i Use `all_of(age_60_vars)` instead of `age_60_vars` to silence this message.
## i See <https://tidyselect.r-lib.org/reference/faq-external-vector.html>.
## This message is displayed once per session.
```

```
head(ny_floodzone_60years)
```

```
## # A tibble: 6 x 7
##   geo_type geo_name geo_id stateusps universe unit          age_60_share
##   <chr>    <chr>    <chr>  <chr>     <chr>    <chr>                <dbl>
## 1 County   Albany   36001  NY        100-Year Single-Family        0.364
## 2 County   Bronx    36005  NY        100-Year Single-Family        0.559
## 3 County   Cayuga   36011  NY        100-Year Single-Family        0.883
## 4 County   Chenango 36017  NY        100-Year Single-Family        0.670
## 5 County   Clinton  36019  NY        100-Year Single-Family        0.663
## 6 County   Cortland 36023  NY        100-Year Single-Family        0.677
```