Floodzone Data Analysis - Census Tract Level

Mattingly, Peter

10 September 2020

Contents

Load Floodzone Data

```
# Import NY Floodzone Data
ny_floodzone <- read_excel("~/Desktop/Thesis Analysis/Data/NY_FloodzoneData.xlsx", sheet = "Census Trac
# Identify "500-Year" and "Combined" Floodplains
ny_floodzone <- ny_floodzone %>%
  rename_all(. %>% tolower %>% gsub(" ", "_", .)) %>%
  mutate(floodplain = case_when(pop_100year == 0 & pop_anyyear > 0 ~ "500-Year",
                                pop_100year > 0 & pop_anyyear > 0 ~ "Combined"))
ny_floodzone %>% count(floodplain)
```

```
## # A tibble: 3 x 2
##   floodplain     n
##   <chr>      <int>
## 1 500-Year     135
## 2 Combined    1043
## 3 <NA>        2864
```

Population

```
pop_vars <- c("geo_type","geo_name","geo_id","stateusps","floodplain",
              "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 7
##   geo_type    geo_name   geo_id    stateusps floodplain universe population
```

```
##   <chr>       <chr>      <chr>     <chr>     <chr>      <chr>         <dbl>
## 1 Census Tra~ 360010001~ 36001000~ NY        Combined   Total          2001
## 2 Census Tra~ 360010002~ 36001000~ NY        <NA>       Total          4519
## 3 Census Tra~ 360010003~ 36001000~ NY        Combined   Total          5244
## 4 Census Tra~ 360010004~ 36001000~ NY        <NA>       Total          2423
## 5 Census Tra~ 360010004~ 36001000~ NY        <NA>       Total          4654
## 6 Census Tra~ 360010005~ 36001000~ NY        <NA>       Total          3758
```

Housing

```
housing_vars <- c("geo_type","geo_name","geo_id","stateusps","floodplain",
                  "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 7
##   geo_type   geo_name  geo_id   stateusps floodplain universe housing_total
##   <chr>      <chr>     <chr>    <chr>     <chr>      <chr>            <dbl>
## 1 Census Tr~ 36001000~ 3600100~ NY        Combined   100-Year         14.7
## 2 Census Tr~ 36001000~ 3600100~ NY        Combined   100-Year          3.54
## 3 Census Tr~ 36001001~ 3600100~ NY        Combined   100-Year          4.09
## 4 Census Tr~ 36001001~ 3600100~ NY        Combined   100-Year         16.1
## 5 Census Tr~ 36001002~ 3600100~ NY        Combined   100-Year        229.
## 6 Census Tr~ 36001002~ 3600100~ NY        Combined   100-Year          8.27
```

Race

```
race_vars <- c("geo_type","geo_name","geo_id","stateusps","floodplain",
               "shr_pop_asian_fp_100","shr_pop_asian_fp_any",
               "shr_pop_asian_tract","shr_pop_black_fp_100","shr_pop_black_fp_any","shr_pop_black_tract
               "shr_pop_hispanic_fp_100","shr_pop_hispanic_fp_any","shr_pop_hispanic_tract","shr_pop_wh
               "shr_pop_white_fp_any","shr_pop_white_tract")

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_tract) %>%
  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 8
##   geo_type  geo_name  geo_id stateusps floodplain race  universe race_share
##   <chr>     <chr>     <chr>  <chr>     <chr>      <chr> <chr>         <dbl>
## 1 Census T~ 36001000~ 36001~ NY        Combined   Asian 100-Year     0.0240
## 2 Census T~ 36001000~ 36001~ NY        Combined   Asian 100-Year     0.0257
## 3 Census T~ 36001001~ 36001~ NY        Combined   Asian 100-Year     0.0796
## 4 Census T~ 36001001~ 36001~ NY        Combined   Asian 100-Year     0.0755
## 5 Census T~ 36001002~ 36001~ NY        Combined   Asian 100-Year     0
## 6 Census T~ 36001002~ 36001~ NY        Combined   Asian 100-Year     0.157
```

Income

```
income_vars <- c("geo_type","geo_name","geo_id","stateusps","floodplain",
                 "shr_hu_incless25_fp_100","shr_hu_incless25_fp_any","shr_hu_incless25_tract","shr_hu_i
                 "shr_hu_inc25to50_fp_any","shr_hu_inc25to50_tract","shr_hu_inc50to75_fp_100","shr_hu_i
                 "shr_hu_inc50to75_tract","shr_hu_inc75up_fp_100","shr_hu_inc75up_fp_any","shr_hu_inc75

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_tract) %>%
  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 8
##   geo_type geo_name geo_id stateusps floodplain income universe
##   <chr>    <chr>    <chr>  <chr>     <chr>      <chr>  <chr>
## 1 Census ~ 3600100~ 36001~ NY        Combined   <$25k  100-Year
## 2 Census ~ 3600100~ 36001~ NY        Combined   <$25k  100-Year
## 3 Census ~ 3600100~ 36001~ NY        Combined   <$25k  100-Year
## 4 Census ~ 3600100~ 36001~ NY        Combined   <$25k  100-Year
## 5 Census ~ 3600100~ 36001~ NY        Combined   <$25k  100-Year
## 6 Census ~ 3600100~ 36001~ NY        Combined   <$25k  100-Year
## # ... with 1 more variable: income_share <dbl>
```

Housing Tenure

```
tenure_vars <- c("geo_type","geo_name","geo_id","stateusps","floodplain",
                 "shr_hu_renter_fp_100","shr_hu_renter_fp_any","shr_hu_renter_tract","shr_hu_owner_fp_1(
                 "shr_hu_owner_fp_any","shr_hu_owner_tract")

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_tract) %>%
  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.
```

```
head(ny_floodzone_tenure)
```

```
## # A tibble: 6 x 8
##   geo_type geo_name geo_id stateusps floodplain tenure universe
##   <chr>    <chr>    <chr>  <chr>     <chr>      <chr>  <chr>
## 1 Census ~ 3600100~ 36001~ NY        Combined   Renter 100-Year
## 2 Census ~ 3600100~ 36001~ NY        Combined   Renter 100-Year
## 3 Census ~ 3600100~ 36001~ NY        Combined   Renter 100-Year
## 4 Census ~ 3600100~ 36001~ NY        Combined   Renter 100-Year
## 5 Census ~ 3600100~ 36001~ NY        Combined   Renter 100-Year
## 6 Census ~ 3600100~ 36001~ NY        Combined   Renter 100-Year
## # ... with 1 more variable: tenure_share <dbl>
```

Building Type

```
unit_vars <- c("geo_type","geo_name","geo_id","stateusps","floodplain",
               "shr_hu_1fam_fp_100","shr_hu_1fam_fp_any","shr_hu_1fam_tract","shr_hu_2to4fam_fp_100",
               "shr_hu_2to4fam_fp_any","shr_hu_2to4fam_tract","shr_hu_5to19fam_fp_100","shr_hu_5to19f
               "shr_hu_5to19fam_tract","shr_hu_20up_fp_100","shr_hu_20up_fp_any","shr_hu_20up_tract",
               "shr_hu_other_fp_100","shr_hu_other_fp_any","shr_hu_other_tract")

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_tract) %>%
  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.
```

```
head(ny_floodzone_unit)
```

```
## # A tibble: 6 x 8
##   geo_type  geo_name  geo_id stateusps floodplain unit  universe unit_share
##   <chr>     <chr>     <chr>  <chr>     <chr>      <chr> <chr>         <dbl>
## 1 Census T~ 36001000~ 36001~ NY        Combined   Sing~ 100-Year     0.306
## 2 Census T~ 36001000~ 36001~ NY        Combined   Sing~ 100-Year     0.309
## 3 Census T~ 36001001~ 36001~ NY        Combined   Sing~ 100-Year     0.0373
## 4 Census T~ 36001001~ 36001~ NY        Combined   Sing~ 100-Year     0.625
## 5 Census T~ 36001002~ 36001~ NY        Combined   Sing~ 100-Year     0.177
## 6 Census T~ 36001002~ 36001~ NY        Combined   Sing~ 100-Year     0.338
```

Poverty

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

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":tract) %>%
  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 7
##   geo_type   geo_name  geo_id   stateusps floodplain universe poverty_share
##   <chr>      <chr>     <chr>    <chr>     <chr>      <chr>            <dbl>
## 1 Census Tr~ 36001000~ 3600100~ NY        Combined   100-Year         0.382
## 2 Census Tr~ 36001000~ 3600100~ NY        Combined   100-Year         0.347
## 3 Census Tr~ 36001001~ 3600100~ NY        Combined   100-Year         0.429
## 4 Census Tr~ 36001001~ 3600100~ NY        Combined   100-Year         0.101
## 5 Census Tr~ 36001002~ 3600100~ NY        Combined   100-Year         0.451
## 6 Census Tr~ 36001002~ 3600100~ NY        Combined   100-Year         0.324
```

Subsidized Housing

```
subsidized_vars <- c("geo_type","geo_name","geo_id","stateusps","floodplain",
                  "ph_units_100year","ph_units_anyyear","sub_housing_units_100year","sub_housing_units_a

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 8
##   geo_type geo_name geo_id stateusps floodplain subsidized universe
##   <chr>    <chr>    <chr>  <chr>     <chr>      <chr>      <chr>
## 1 Census ~ 3600100~ 36001~ NY        Combined   Public Ho~ 100-Year
## 2 Census ~ 3604700~ 36047~ NY        Combined   Public Ho~ 100-Year
## 3 Census ~ 3604703~ 36047~ NY        Combined   Public Ho~ 100-Year
## 4 Census ~ 3604703~ 36047~ NY        Combined   Public Ho~ 100-Year
## 5 Census ~ 3605941~ 36059~ NY        Combined   Public Ho~ 100-Year
## 6 Census ~ 3605941~ 36059~ NY        Combined   Public Ho~ 100-Year
## # ... with 1 more variable: subsidized_units <dbl>
```

Building Age

```
years_vars <- c("geo_type","geo_name","geo_id","stateusps","floodplain",
                "shr_hu_builtpre60_fp_100","shr_hu_builtpre60_fp_any","shr_hu_builtpre60_tract","shr_hu_bu
                "shr_hu_built6079_fp_any","shr_hu_built6079_tract","shr_hu_built8099_fp_100","shr_hu_built
                "shr_hu_built8099_tract","shr_hu_built00s_fp_100","shr_hu_built00s_fp_any","shr_hu_built0(

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_tract) %>%
  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 8
##   geo_type geo_name geo_id stateusps floodplain housing_age universe
##   <chr>    <chr>    <chr>  <chr>     <chr>      <chr>       <chr>
## 1 Census ~ 3600100~ 36001~ NY        Combined   60+         100-Year
## 2 Census ~ 3600100~ 36001~ NY        Combined   60+         100-Year
## 3 Census ~ 3600100~ 36001~ NY        Combined   60+         100-Year
## 4 Census ~ 3600100~ 36001~ NY        Combined   60+         100-Year
## 5 Census ~ 3600100~ 36001~ NY        Combined   60+         100-Year
## 6 Census ~ 3600100~ 36001~ NY        Combined   60+         100-Year
## # ... with 1 more variable: age_share <dbl>
```

Buildings 60+ Years

```
age_60_vars <- c("geo_type","geo_name","geo_id","stateusps","floodplain",
                 "shr_hu_pre60_fp_100_1fam","shr_hu_pre60_fp_any_1fam","shr_hu_pre60_fp_100_2to4fam","shr_
                 "shr_hu_pre60_fp_100_5to19fam","shr_hu_pre60_fp_any_5to19fam","shr_hu_pre60_fp_100_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 8
##   geo_type geo_name geo_id stateusps floodplain universe unit  age_60_share
##   <chr>    <chr>    <chr>  <chr>     <chr>      <chr>    <chr>        <dbl>
## 1 Census ~ 3600100~ 36001~ NY        Combined   100-Year Sing~       0.321
## 2 Census ~ 3600100~ 36001~ NY        Combined   100-Year Sing~       0.315
## 3 Census ~ 3600100~ 36001~ NY        Combined   100-Year Sing~       0.0723
## 4 Census ~ 3600100~ 36001~ NY        Combined   100-Year Sing~       0.750
## 5 Census ~ 3600100~ 36001~ NY        Combined   100-Year Sing~       0.188
## 6 Census ~ 3600100~ 36001~ NY        Combined   100-Year Sing~       0.407
```