

# Instructions for Adding Council District (CD) and Police Precinct (PP) Data to 311 Service Requests

## Step 0: Downloading and Formatting Data for 2010-2024

Before proceeding, ensure that you have downloaded and formatted the 311 service request data for all years between 2010 and 2024. Follow the instructions provided in the **Instructions\_downloading\_311\_data.pdf** file to complete this step. You should have 15 RDS files, one for each year, saved on your computer.

## Step 1: Downloading the Shapefiles

1. **Visit the NYC Planning Open Data Portal:**
  - Go to the [NYC Planning Open Data Portal](#).
2. **Download the Shapefiles:**
  - Scroll down to the section titled "Political and Elections Districts."
  - Download the shapefile for **City Council Districts (Water Areas Included)**.
  - Next, scroll to the section titled "School, Police, Health and Fire."
  - Download the shapefile for **Police Precincts (Clipped to Shoreline)**.
  - Save these files to a known location on your computer.
3. **Extract the Files:**
  - Ensure that the downloaded shapefiles are extracted and that all associated files (.shp, .shx, .dbf, etc.) are in the same directory.

## Step 2: Adding Council District and Police Precinct Data

1. **Open RStudio:**
  - Start RStudio and open a new R script file.
2. **Copy-Paste the Script Below:**
  - Copy the following script into your R script file.

```

library(data.table)
library(sf)

# Set file paths
rds_file <- "path/to/rds/file/SRs_2010.rds"
council_districts_shapefile <- "/path/to/CD/shapefile/nyccwi.shp"
police_precincts_shapefile <- "/path/to/PP/shapefile/nypp.shp"
output_file <- "/path/to/output/file/SRs_2010_with_CD_PP.rds"

# Load the shapefiles for council districts and police precincts
council_districts <- st_read(council_districts_shapefile)
police_precincts <- st_read(police_precincts_shapefile)

# Transform the shapefiles to the same CRS as the latitude and longitude
data
council_districts <- st_transform(council_districts, crs = 4326)
police_precincts <- st_transform(police_precincts, crs = 4326)

# Function to add council district and police precinct based on latitude
and longitude
add_spatial_info <- function(data, council_districts, police_precincts) {
  # Ensure that longitude and latitude columns exist
  if (!("Longitude" %in% names(data)) | !("Latitude" %in% names(data))) {
    stop("Longitude and Latitude columns are required in the data.")
  }

  # Initialize the new columns with NA for rows without lat/long info
  data[, council_district := NA_integer_]
  data[, police_precinct := NA_integer_]

  # Filter only the rows with latitude and longitude information for
  spatial joins
  valid_rows <- data[!is.na(Longitude) & !is.na(Latitude)]

  # Convert the valid rows to an sf object
  data_sf <- st_as_sf(valid_rows, coords = c("Longitude", "Latitude"), crs
= 4326, remove = FALSE)

  # Perform spatial joins
  data_with_cd <- st_join(data_sf, council_districts, join = st_intersects)
  data_with_cd_pp <- st_join(data_with_cd, police_precincts, join =
st_intersects)

```

```

# Convert back to data.table and remove geometry column
data_with_cd_pp_dt <- as.data.table(st_drop_geometry(data_with_cd_pp))

# Retain only the necessary columns and add council district and police
precinct
data_with_cd_pp_dt <- data_with_cd_pp_dt[, .(Longitude, Latitude,
CounDist, Precinct)]
setnames(data_with_cd_pp_dt, c("CounDist", "Precinct"),
c("council_district", "police_precinct"), skip_absent = TRUE)

# Merge the updated columns back into the original dataset
data[data_with_cd_pp_dt, on = .(Longitude, Latitude),
:=`(council_district = i.council_district, police_precinct =
i.police_precinct)]

return(data)
}

# Load the RDS file
SRs_2010 <- readRDS(rds_file)

# Split data into smaller batches to process
batch_size <- 500000 # Adjust this size based on memory capacity
num_batches <- ceiling(nrow(SRs_2010) / batch_size)

for (i in 1:num_batches) {
  start_index <- (i - 1) * batch_size + 1
  end_index <- min(i * batch_size, nrow(SRs_2010))
  SRs_batch <- SRs_2010[start_index:end_index]

  # Add spatial information to the batch
  SRs_batch <- add_spatial_info(SRs_batch, council_districts,
police_precincts)

  # Save each batch to a file or combine them
  if (i == 1) {
    saveRDS(SRs_batch, output_file)
  } else {
    # Append to the output file
    existing_data <- readRDS(output_file)
    combined_data <- rbind(existing_data, SRs_batch)
    saveRDS(combined_data, output_file)
  }
}

```

```
rm(SRs_batch) # Clean up memory
gc() # Run garbage collection
}
rm(SRs_2010, council_districts, police_precincts)
gc()
```

### 3. Modify the File Paths:

- Replace "`path/to/your/SRs_2010.rds`" with the actual path to your existing RDS file for the year 2010.
- Replace "`path/to/your/City_Council_Districts.shp`" and "`path/to/your/Police_Precincts.shp`" with the paths to your downloaded shapefiles.
- Replace "`path/to/your/SRs_2010_with_CD_PP.rds`" with the path to your output rds file.

### 4. Run the Script:

- Run the entire script in RStudio. This will:
  - Load your existing 311 service requests data.
  - Add Council District and Police Precinct information based on the geographic coordinates.
  - Save the updated dataset with CD and PP information to a new RDS file.

### 5. Verify the Updated Data:

- Check the newly created RDS file to ensure that it now contains the `council_district` and `police_precinct` columns.

## Step 3: Repeat for Other Years

You need to repeat the process outlined in Step 2 for each year from 2010 to 2024. Here's how you can do it:

#### 1. Adjust the File Paths:

- For each year, replace the `rds_file` path with the path to the corresponding RDS file for that year.
- Modify the `output_rds_file` path to ensure the updated file is saved with a distinct name (e.g., "`path/to/save/SRs_2011_with_CD_PP.rds`" for the year 2011).

#### 2. Run the Script:

- Once the file paths have been updated, re-run the script for each year.

- This will add Council District and Police Precinct information to the data for that specific year.
3. **Verify and Save:**
    - After running the script, verify that the `council_district` and `police_precinct` columns are present in each of the newly created RDS files.
  4. **Continue for All Years:**
    - Repeat this process for each year from 2010 to 2024 until you have updated and saved the datasets for all 15 years.