Relevant Data for Fire Project

5/28/2020

Summary

The following is a R markdown file that identifies the data tables we need from the Charlottesville open data portal (https://opendata.charlottesville.org). Additionally, it provides R code to pull the relevant features from each of the tables. The data were downloaded between April 23rd and May 26th of 2020, so versions of these tables downloaded outside of this range may necessitate modifications of certain variable names.

Needed

- Real Estate (Base Data)
- Real Estate (Residential Details)
- Real Estate (Commerical Details)
- Parcel Area Details
- Parcel Owner Points
- Building Permits Spatial
- Existing Structure Area

May Need

- Parcel Boundary Area
- Master Address Points
- Master Address Table

Pulling Relevant Features

Real Estate:

```
RE_base = read.csv('Real_Estate_Base_Data.csv')
to_drop = c("RecordID_Int",
            "TaxType",
            "TaxDist",
            "Legal")
RE_base = RE_base[,!(names(RE_base) %in% to_drop)]
c("REAL ESTATE (BASE DATA) FEATURES:",names(RE_base))
```

```
## [1] "REAL ESTATE (BASE DATA) FEATURES:" "ParcelNumber"
## [3] "StreetNumber"                      "StreetName"
## [5] "Unit"                              "StateCode"
## [7] "Zone"                              "Acreage"
## [9] "GPIN"
```

```
RE_resid = read.csv('Real_Estate_Residential_Details.csv')
to_drop = c("RecordID_Int",
            "Flooring",
            "HalfBathrooms")
RE_resid = RE_resid[,!(names(RE_resid) %in% to_drop)]
c("REAL ESTATE (RESIDENTIAL DETAILS) FEATURES:",names(RE_resid))
```

```
##  [1] "REAL ESTATE (RESIDENTIAL DETAILS) FEATURES:"
##  [2] "ParcelNumber"
##  [3] "StreetNumber"
##  [4] "StreetName"
##  [5] "Unit"
##  [6] "UseCode"
##  [7] "Style"
##  [8] "Grade"
##  [9] "Roof"
## [10] "Heating"
## [11] "Fireplace"
## [12] "YearBuilt"
## [13] "TotalRooms"
## [14] "Bedrooms"
## [15] "FullBathrooms"
## [16] "BasementGarage"
## [17] "Basement"
## [18] "FinishedBasement"
## [19] "BasementType"
## [20] "ExternalWalls"
## [21] "NumberOfStories"
## [22] "SquareFootageFinishedLiving"
```

```
RE_comm = read.csv('Real_Estate_Commercial_Details.csv')
to_drop = c("RecordID_Int")
RE_comm = RE_comm[,!(names(RE_comm) %in% to_drop)]
c("REAL ESTATE (COMMERCIAL DETAILS) FEATURES:",names(RE_comm))
```

```
##  [1] "REAL ESTATE (COMMERCIAL DETAILS) FEATURES:"
##  [2] "ParcelNumber"
##  [3] "UseCode"
##  [4] "YearBuilt"
##  [5] "GrossArea"
##  [6] "StoryHeight"
##  [7] "NumberOfStories"
##  [8] "StreetName"
##  [9] "StreetNumber"
## [10] "Unit"
```

Parcel:

```
PR_area = read.csv('Parcel_Area_Details.csv')
to_drop = c("OBJECTID",
            "IsMultiParcelPolygon",
            "Label",
            "LegalDescription",
            "MapPage",
            "ModifiedDate",
            "OwnerName",
            "OwnerAddress",
            "OwnerCityState",
            "OwnerZipCode",
            "Text",
            "ESRI_OID")
PR_area = PR_area[,!(names(PR_area) %in% to_drop)]
c("PARCEL AREA DETAILS FEATURES:",names(PR_area))
```

```
##  [1] "PARCEL AREA DETAILS FEATURES:" "Assessment"
##  [3] "FileType"                      "GeoParcelIdentificationNumber"
##  [5] "LotSquareFeet"                 "ParcelNumber"
##  [7] "StreetName"                    "StreetNumber"
##  [9] "TaxYear"                       "Unit"
## [11] "Zoning"
```

```
PR_points = read.csv('Parcel_Owner_Points.csv')
to_drop = c("OBJECTID",
            "FileType",
            "LegalDescription",
            "LotSquareFeet",
            "ParcelMap",
            "UseCode")
PR_points = PR_points[,!(names(PR_points) %in% to_drop)]
c("PARCEL OWNER POINTS FEATURES:",names(PR_points))
```

```
##  [1] "PARCEL OWNER POINTS FEATURES:" "X"
##  [3] "Y"                             "GeoParcelIdentificationNumber"
##  [5] "OwnerName"                     "ParcelNumber"
##  [7] "StreetName"                    "StreetNumber"
##  [9] "Unit"                          "ZipCode"
## [11] "Zone"
```

Building Permits and Exisiting Structure:

```
Building_Permits = read.csv('Building_Permits_Spatial_.csv')
to_keep = c("IssuedDate", "ParcelNumber", "SubType", "Type", "WorkDescription")
Building_Permits = Building_Permits[,c(to_keep)]
c("BUILDING PERMITS SPATIAL FEATURES:",names(Building_Permits))
```

```
## [1] "BUILDING PERMITS SPATIAL FEATURES:" "IssuedDate"
## [3] "ParcelNumber"                       "SubType"
## [5] "Type"                               "WorkDescription"
```

```
Existing_Structure = read.csv('Existing_Structure_Area.csv')
to_keep = c("BIN", "ST_NUMBER", "STREET")
Existing_Structure = Existing_Structure[,c(to_keep)]
c("EXISTING STRUCTURE AREA FEATURES:",names(Existing_Structure))
```

```
## [1] "EXISTING STRUCTURE AREA FEATURES:" "BIN"
## [3] "ST_NUMBER"                         "STREET"
```

Features of the "May Need" Data

```
PR_bound = read.csv('Parcel_Boundary_Area.csv')
c("PARCEL BOUNDARY AREA FEATURES:",names(PR_bound))
```

```
## [1] "PARCEL BOUNDARY AREA FEATURES:" "GPIN"
## [3] "OBJECTID"
```

```
Master_Address_Points = read.csv('Master_Address_Points.csv')
c("MASTER ADDRESS POINTS FEATURES:",names(Master_Address_Points))
```

```
##  [1] "MASTER ADDRESS POINTS FEATURES:" "X"
##  [3] "Y"                               "OBJECTID"
##  [5] "geo_MAT_Status"                  "geo_MAT_ST_NUMBER"
##  [7] "geo_MAT_PREDIR"                  "geo_MAT_ST_NAME"
##  [9] "geo_MAT_SUFFIX"                  "geo_MAT_POSTDIR"
## [11] "centroid_BUILD_STREET"           "unit"
## [13] "address"                         "geo_MAT_MasterAddressID"
```

```
Master_Address_Table = read.csv('Master_Address_Table.csv')
c("MASTER ADDRESS TABLE FEATURES:",names(Master_Address_Table))
```

```
##  [1] "MASTER ADDRESS TABLE FEATURES:" "OBJECTID"
##  [3] "BIN"                            "ST_NUMBER"
##  [5] "ST_UNIT"                        "UNIT_TYPE"
##  [7] "PREDIR"                         "ST_NAME"
##  [9] "SUFFIX"                         "POSTDIR"
## [11] "ZIP"                            "STSEGID"
## [13] "USE_TYPE"                       "MasterAddressID"
```