

Investigating the 2021 Texas Power Crisis

EDS-223 Assignment 3

Isabella Segarra

2025-11-10

Texas Power Crisis

texas-power-crisis

This repository's purpose is for investigating the 2021 Texas power crisis caused by extreme weather events.

Investigating the 2021 Texas Power Crisis

Contributor:

Isabella Segarra (Masters of Environmental Data Science, UC Santa Barbara)

Contents:

This repository's purpose is for investigating the Texas Power Crisis that took place in February 2021 after a series of three severe storms.

It contains the following file structure:



How to access the data:

The night lights data is from [NASA's Leve-1 and Atmospheric Archive and Distribution System Distributed Active Archive Center (LAADS DAAC)] (<https://ladsweb.modaps.eosdis.nasa.gov/>). The roads and house/building data is from [OpenStreetMap (OSM)] (<https://planet.openstreetmap.org/>), an open-source collaborative project for geographic data. The socioeconomic data is from the [U.S. Census American Community Survey] (<https://www.census.gov/programs-surveys/acs>)

Citations:

Román, M.O., Wang, Z., Sun, Q., Kalb, V., Miller, S.D., Molthan, A., Schultz, L., Bell, J., Stokes, E.C., Pandey, B. and Seto, K.C., et al. (2018). NASA's Black Marble nighttime lights product suite. *Remote Sensing of Environment* 210, 113-143. doi:10.1016/j.rse.2018.03.017.

OpenStreetMap. (n.d.) Planet OSM. OpenStreetMap. (<https://planet.openstreetmap.org/>)

Geofabrik GmbH. (n.d.). OpenStreetMap Data Extracts. GEOFABRIK. (<https://download.geofabrik.de/>)

United States Census Bureau. 2019. "American Community Survey (ACS)." (<https://www.census.gov/programs-surveys/acs>)

Acknowledgements:

I would like to acknowledge Dr. Annie Adams and Ale Vidal Meza for their contributions to this project.

Figure 1:³ README

Objective

In this project, I investigated the February 2021 Texas power crisis, which was triggered by three severe winter storms that caused widespread blackouts across the region. To illustrate the impact of the storms, I mapped Houston's night lights before and after the crisis and identified areas that experienced a blackout. To explore the differential effects of the power outages from a socioeconomic perspective, I compared the median household incomes of census tracts that experienced a blackout with those that did not.

Environment Set-up

```
# Load relevant libraries
library(sf) # For vector data
library(stars) # For raster data
library(terra) # raster handling
library(tmap) # For static and interactive maps
library(here) # For importing data
library(tidyverse) # For data cleaning
library(dplyr) # For filtering data
library(paletteer) # For pretty colors
library(testthat) # For efficient workflows
library(kableExtra) # For pretty tables
library(patchwork) # For plotting
```

Data Import

The data I am using in this project includes data from NASA's Worldview for night light data, data for roads and houses from OpenStreetMap (OSM), and socioeconomic data from the U.S. Census Bureau's American Community Survey.

Night lights data import

This raster data is from NASA's Worldview tool, distributed through NASA's Level-1 and Atmospheric Archive & Distribution System Distributed Active Archive Center (LAADS DAAC). It is a .tif file that contains 10x10 degree tiles. I will be importing four data files from the VNP46A1 folder for night light data from the days surrounding the storm as stars.

```

# Import tile h08v05 from 2021-02-07
nl_h08v05_07 <- read_stars(here("data", "VNP46A1","VNP46A1.A2021038.h08v05.001.2021039064328

# Import tile h08v06 from 2021-02-07
nl_h08v06_07 <- read_stars(here("data", "VNP46A1","VNP46A1.A2021038.h08v06.001.2021039064329

# Import tile h08v05 from 2021-02-16
nl_h08v05_16 <- read_stars(here("data", "VNP46A1","VNP46A1.A2021047.h08v05.001.2021048091106

# Import tile h08v06 from 2021-02-16
nl_h08v06_16 <- read_stars(here("data", "VNP46A1","VNP46A1.A2021047.h08v06.001.2021048091105

```

Roads data import

This vector data is from OSM, processed with the help of a third party site to produce a shapefile of all Texas highways. This data contains 6,085 observations with 11 variables.

```

# Import roads data with SQL query to filter for only highways (use st_read)
roads <- read_sf(here("data", "gis_osm_roads_free_1.gpkg"), query = "SELECT * FROM gis_osm_r

```

House data import

This vector data is also from OSM, processed with the help of a third party to produce a shapefile of houses in the Houston metropolitan area. This data contains 475941 observations with 6 variables.

```

# Import house data with SQL query to filter for selected building types (use st_read)
houses <- read_sf(here("data", "gis_osm_buildings_a_free_1.gpkg"), query = "SELECT * FROM gis_g

```

Socioeconomic data import

This data is from U.S. Census Bureau's American Community Survey 2019 census. This Geodatabase (gdb) contains data for Texas. I will be importing two layers from the gdb.

```

# Import socioeconomic geometry layer
socioeco_geom <- read_sf(here("data", "ACS_2019_5YR_TRACT_48_TEXAS.gdb"), layer = "ACS_2019_5YR_TRACT_48_TEXAS

# Import socioeconomic income layer
socioeco_income <- read_sf(here("data", "ACS_2019_5YR_TRACT_48_TEXAS.gdb"), layer = "X19_INCOME

```

```
# View layers with 'st_layers' (commented out for now)
#st_layers(here("data", "ACS_2019_5YR_TRACT_48_TEXAS.gdb"))
```

Data Exploration

CRS Matching

Before diving into the data, I want to see what CRSs the data are currently set to.

```
# .....Check CRS of individual datasets.....
# Create list of the datasets that have geometries
datasets <- list(houses, roads, socioeco_geom, nl_h08v05_07, nl_h08v06_07, nl_h08v05_16, nl_h08v06_16)

# Loop through each dataset and print its CRS

for (i in seq_along(datasets)) {
  epsg <- st_crs(datasets[[i]])$epsg
  print(epsg)
}
```

Check for Invalid geometries

Check for invalid geometries in our three datasets (roads, houses, socioeco_geom) that contain a geometry type (e.g, multipolygon).

```
#.....Checking for invalid geometries.....
# Create a dataset with only the data that contains geometries
datasets_geom = list(roads, houses, socioeco_geom)

# Check if there are invalid geometries
for (i in seq_along(datasets_geom)) {
  which(!st_is_valid(datasets_geom[[i]]))
  message("There are NO invalid geometries")
}
```

Night Lights

Merge tiles

In this section, I merged the tiles for February 7th, 2021 and February 16, 2021.

```
#.....Night light change .....
```

```
# Merge night light for 2021-02-07
nl_07 <- st_mosaic(nl_h08v05_07, nl_h08v06_07)

# Merge night light for 2021-02-16
nl_16 <- st_mosaic(nl_h08v05_16, nl_h08v06_16)

# Confirm class is
print(class(nl_07))
print(class(nl_16))

# Clear up environment for efficiency
rm(list = c("nl_h08v05_07", "nl_h08v06_07"))
rm(list = c("nl_h08v05_16", "nl_h08v06_16"))
```

Map: Night Lights

In this section, I mapped the night lights for February 7th and 16th. In order to do so, I created a bounding box for the Houston area to crop the night light data with.

```
# Create bounding box for Houston area
bbox_houston <- st_bbox(c(xmin = -96.5, xmax = -94.5, ymax = 30.5, ymin = 29), crs = st_crs(ni...))

# Crop Feb 7th night light data
nl_07_crop<- st_crop(nl_07, bbox_houston)

# Crop Feb 16th night light data
nl_16_crop<- st_crop(nl_16, bbox_houston)

# Verify that 'nl_07' and 'nl_16' were cropped to Houston area

if(ncell(nl_07) == nrow(nl_07_crop)) {
  warning("data did not crop!")
} else {
  message("data cropped!")
```

```

}

if(ncell(nl_16) == nrow(nl_16_crop)) {
  warning("data did not crop!")
} else {
  message("data cropped!")
}

#.....Mapping 2021-02-07.....  

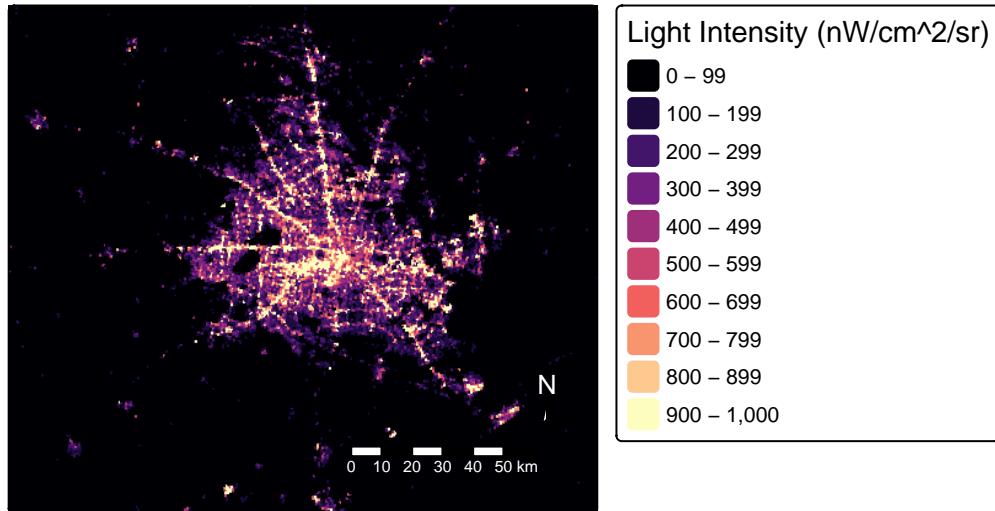
nl_07_map <- tm_shape(nl_07_crop) + # Map cropped data
  tm_raster(palette = "magma",
             breaks = c(0, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000), # custom breaks
             title = "Light Intensity (nW/cm^2/sr)") +
  tm_title("Houston, Texas Night Lights", fontface = "bold") +
  tm_title("Before storm: 02-07-2021", size = 0.8) +
  tm_layout(bg.color = "black", legend.show = TRUE) +
  tm_compass(size = 0.8, text.color = "white",
             color.light = "white") +
  tm_scalebar(size = 0.5, text.color = "white",
              color.light = "white")

# View map
nl_07_map

```

Houston, Texas Night Lights

Before storm: 02-07-2021



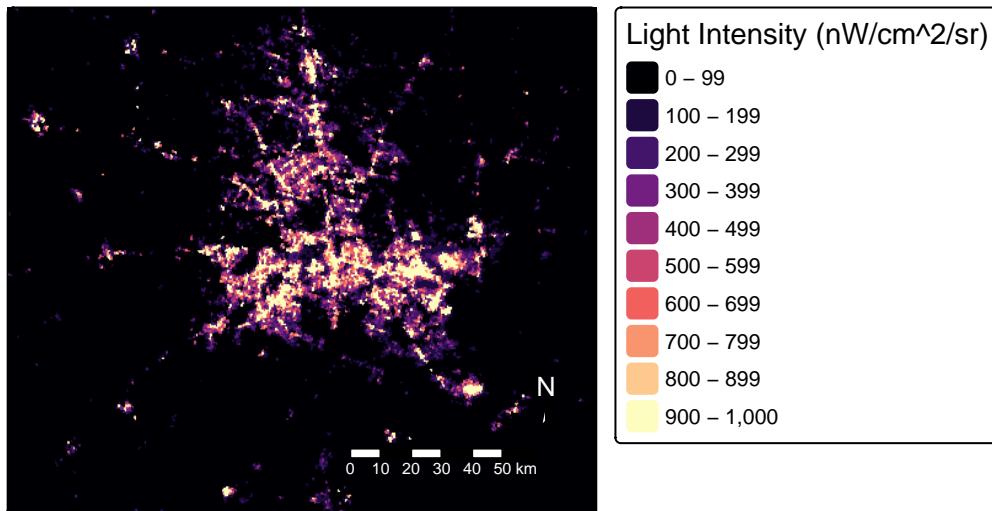
```
#.....Mapping 2021-02-16.....
```

```
nl_16_map <- tm_shape(nl_16_crop) + # map cropped data
  tm_raster(palette = "magma",
             breaks = c(0, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000), # custom breaks
             title = "Light Intensity (nW/cm^2/sr)") +
  tm_title("Houston, Texas Night Lights", fontface = "bold") +
  tm_title("After storm: 02-16-2021 ", size = 0.8) +
  tm_layout(bg.color = "black", legend.show = TRUE) +
  tm_compass(size = 0.8, text.color = "white",
             color.light = "white") +
  tm_scalebar(size = 0.5, text.color = "white",
              color.light = "white")

# View map
nl_16_map
```

Houston, Texas Night Lights

After storm: 02-16-2021



Find blackouts

Find blackout regions by defining the blackout regions where more than 200 nW cm-2sr-1 experienced a blackout and those less than 200 nW cm-2sr-1 did not experience a blackout.

```
#.....Vector difference mask .....
```

```
# Create conditional where TRUEs are regions that experienced a blackout
blackout <- (nl_16 - nl_07) > 200

# Reassign FALSEs as NA
blackout[blackout == FALSE] <- NA

# Turn into vector object
blackout <- st_as_sf(blackout)

# View your new vector dataframe
#head(blackout)

# Check for invalid geometries
if (any(!st_is_valid(blackout))) {
  blackout <- st_make_valid(blackout)
```

```

} else if (nrow(blackout) == 0) {
  print("No invalid geometries found.")
} else {
  print("All geometries valid!")
}

#.....Crop 'Blackout' mask to Houston area ......

# Crop your sf object to Houston bbox made previously
blackout <- st_crop(blackout, bbox_houston)

# Re-project to EPSG:3083 (NAD83 / Texas Centric Albers Equal Area)
blackout <- st_transform(blackout, crs = "EPSG:3083")

# Test that CRS re-projected
test_that("Test 1: CRS", {
  expect_equal(st_crs(blackout)$epsg, 3083)
})

```

Test passed

Highways

In this section, I will exclude highways from the blackout mask because they may have experienced night light changes not associated with the severe storms.

```

#.....Create buffer of 200 meters ......

# The CRS of 'roads' needs to be transformed to the Texas CRS
roads <- st_transform(roads, crs = "EPSG:3083")

# Check the units of 'roads'
units_roads <- st_crs(roads)$units

if (units_roads != "m") {
  warning("Units are NOT in meters!")
} else {
  message("Units are in meters!")
}

```

```

# Identify areas within 200 m of all highways
highway <- st_buffer(roads, dist = 200) %>%
  st_union()

# Check for invalid geometries
which(!st_is_valid(highway))

# Check that buffer is a polygon
st_geometry_type(highway)

#.....Areas further than 200 m.....
blackout_highways <- st_difference(blackout, highway)

```

Buildings

In this section I found the buildings in Houston, TX affected by the blackout. I used the function ‘st_join’ by intersection to find the houses all within the ‘blackout_highways’ multipolygon data frame that experienced a blackout.

```

#.....Find buildings affected by blackout ......

# The CRS of 'roads' needs to be transformed to the Texas CRS
houses <- st_transform(houses, crs = "EPSG:3083")

# Join blackout highways and houses
buildings_blackout <- st_join(houses, blackout_highways, join = st_intersects, left = TRUE)

# Update column name to blackout
buildings_blackout<- buildings_blackout %>% rename(blackout = VNP46A1.A2021047.h08v05.001.2021047)

# View the class of the new data frame
print(class(buildings_blackout))

```

For the estimate of number of homes that experienced blackouts, I filtered the buildings that experienced a blackout to just those of type house, apartments, and residential.

```

# .....Estimate number of houses that experienced blackout ......

# Filter buildings_blackout for only select buildings
residential_blackout <- buildings_blackout %>%

```

```

filter(type %in% c("house", "apartments", "residential"))

# Estimate the sum of blackouts per building type
residential_estimate <- residential_blackout %>%
  group_by(type) %>%
  summarize(blackout_buildings = sum(blackout, na.rm = TRUE))

```

Table 1: Table 1. Estimate of houses that experienced blackouts

Type of building	Number of Blackouts
apartments	609
house	6275
residential	278

```

#.....Find buildings not affected by blackout.....
buildings_no_blackout <- buildings_blackout %>% #running slow, dropping geom
  filter(is.na(blackout))

```

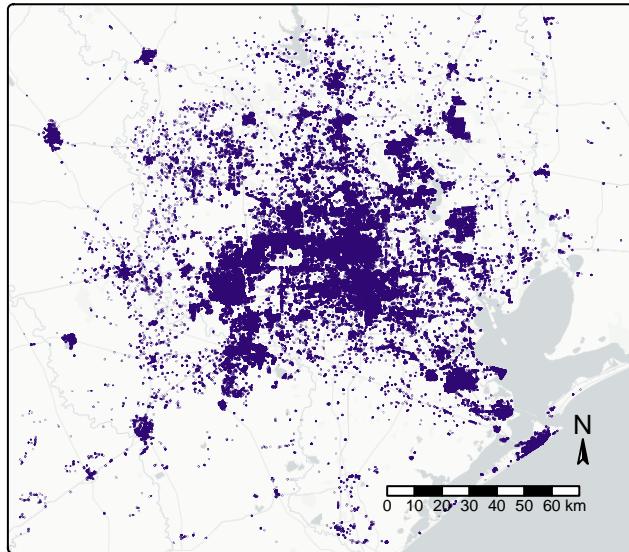
Map: Homes in Houston that lost power

```

tm_shape(buildings_blackout) +
  tm_polygons(col = "#2f0873") +
  tm_title("Houses in Houston Affected by the 2021 Power Crisis", fontface = "bold") +
  tm_layout(bg.color = "white") +
  tm_basemap("CartoDB.PositronNoLabels") +
  tm_compass(size = 0.8) +
  tm_scalebar(text.size = 0.5)

```

Houses in Houston Affected by the 2021 Power Crisis



Census Tracts

In this section I investigated which census tracts experienced blackouts or no blackouts.

Some important information for this analysis:

- socioeco_geom = census tracts geometries
- socioeco_income = median income
- B19013m1 = Median household income in the past 12 months

```
#.....Data wrangling.....  
  
# rename GEOID column in socioeco_geom  
socioeco_geom <- socioeco_geom %>%  
  select(-GEOID) %>%  
  rename(GEOID = GEOID_Data)  
  
# Join layers of socioeco and select columns relevant for analysis  
socieoco <- left_join(socioeco_geom, socioeco_income, by = "GEOID") %>%  
  select(GEOID, B19013m1, Shape) %>%  
  rename(geom_soc = Shape)  
  
# Check CRS of 'socieoco' and 'buildings_blackout'  
if(st_crs(socieoco) == st_crs(buildings_blackout)) {  
  print("coordinate reference systems match")}
```

```

} else{
  print("coordinate reference systems do NOT match")
}

# Transform CRS
if(st_crs(socieoco)$epsg != 3083) {
  print("CRS not for Texas! Transforming...")
  socieoco <- st_transform(socieoco, crs = "EPSG:3083")
}

#.....Census tracts that experienced blackout ......

soc_blackout <- st_join(socieoco, buildings_blackout, left = TRUE)

# Rename column with median incomes to 'income_blackout'
soc_blackout <- soc_blackout %>% rename(income_blackout = B19013m1)

#.....Census tracts that did NOT experience blackout ......

soc_no_blackout <- st_join(socieoco, buildings_no_blackout, left = TRUE)

# Rename column with median incomes to 'income_no_blackout'
soc_no_blackout <- soc_no_blackout %>% rename(income_no_blackout = B19013m1)

#..... Rejoin datasets ......

# Rejoin data set. Error: vector memory limit reached.
#soc_join <- soc_blackout %>% st_drop_geometry() %>%
#  #left_join(soc_no_blackout %>% st_drop_geometry(), by = "GEOID")

# make box plot of median income with 'income_blackout' and 'income_no_blackout' data

# plot of median income blackout regions
library(patchwork)

blackout_soc <- ggplot(soc_blackout, aes(x = income_blackout)) +
  geom_histogram(fill = "#b4c29d") +
  theme_bw() +
  labs(x = "Median Income",
       title = "Homes Affected by Blackouts")

```

```

no_blackout_soc <- ggplot(soc_no_blackout, aes(x = income_no_blackout)) +
  geom_histogram(fill = "#f58056") +
  theme_bw() +
  labs(x = "Median Income",
       title = "Homes Not Affected by Blackouts")

blackout_soc + no_blackout_soc

```

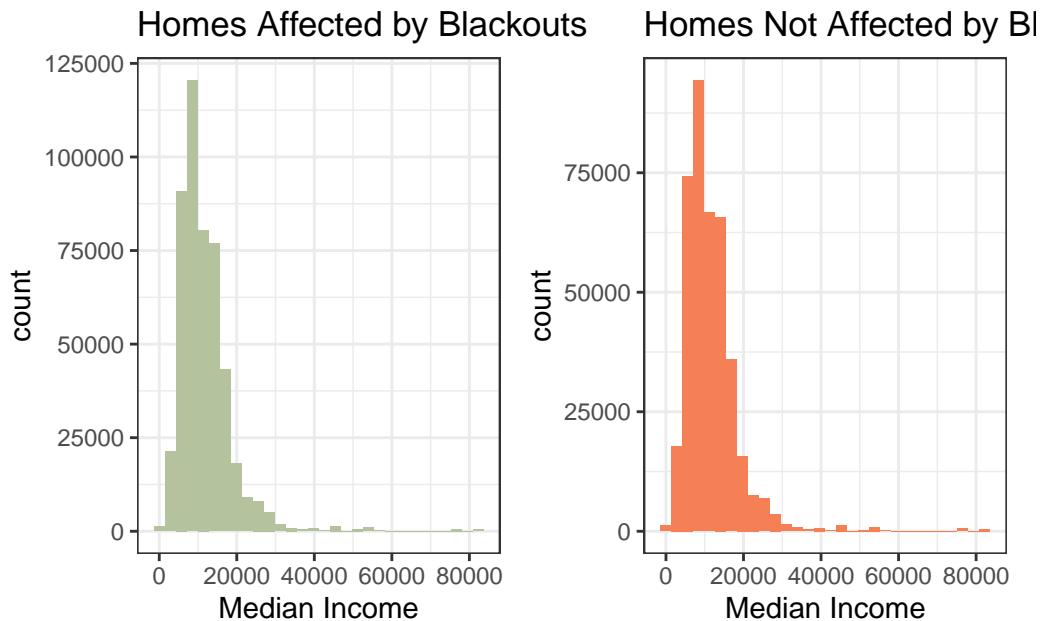


Figure 2: Distribution of median incomes for homes affected and not affected by blackouts.

Map of Census Tracts that lost power

Below are attempts at mapping the census tracts that lost power. Tmap() is struggling to map this large of a dataset. I plan to continue to troubleshoot this error.

```

# Attempts to fix issue

#class(soc_blackout)
#nrow(soc_blackout)
#st_geometry(soc_blackout)

```

```
# Simplify the data
#soc_blackout_simple <- st_simplify(soc_blackout, dTolerance = 100)
# Filter for only relevant columns (fail to run)
#soc_blackout_simple <- soc_blackout_simple %>%
  #select(GEOID, income_blackout)

#tm_shape(soc_blackout) %>%
  #tm_polygons()
```

Discussion

Based on the results of this investigation, it is clear that the 2021 Texas power crisis had widespread impacts across the state and even more pronounced effects in the city of Houston. From observing changes in night lights, I saw that large regions of Houston experienced major blackouts spanning more than 10km. Looking at Census tract data, more houses with median incomes below \$20,000 per year experienced blackouts compared to those that did not. As with many dense urban areas, much of the variability in the data is not captured. Because of this, it is safe to assume that the effects on Texan residents—especially in cities like Houston—are likely underrepresented.