Impact of the High Population Events on Urban Mobility: Who is Affected the Most?

Jane Roarty, Kieu Giang Nguyen, Son Phan

December 09, 2019

Abstract

This project explores mobility in Washington, D.C. in normal conditions and during highly populated events that create a shock to the transportation system. Washington is significant as the nation's capital and therefore a host of major events. Washington is also a considerably segregated city. Because of the frequent events and segregation, it is important to model mobility throughout the city and the metropolitan area. Specifically, this paper models normal traffic conditions in Washington and conditions during the 2017 inauguration and subsequent Women's March by using Spatial Autoregressive Models.

Introduction

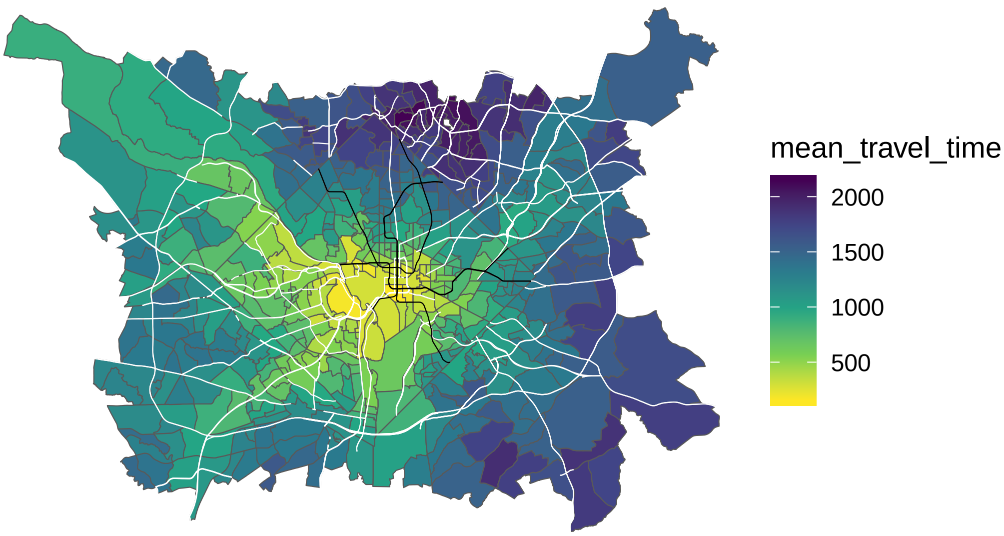
This paper explores mobility in Washington, D.C. during major events such as the 2017 women's march and inauguration. On January 20, 2017, Donald Trump was inaugurated in Washington. The following day, January 21, hundreds of thousands of people protested Trump's inauguration through the Women's March. The women's march in Washington has been called the largest protest in the United States with an estimated 470,000 people in attendance (Wallace and Parlapiano 2017). Crowd scientists estimate that the crowd at the Trump inauguration was about a third of the size, making it around 160,000 people (Wallace and Parlapiano 2017). Population increases of this size over the course of just one weekend creates a huge shock to the transportation system. Before modeling this impact, however, we first modeled the impact of demographic census data on travel time during normal January conditions. We then were able to observe how demographics impact mobility and how those factors are exacerbated during major shocks to the system. We found that the number of roads, distance to the Washington Monument, and the number of black residents are consistently the most significant predictors of travel time to each polygon from the Washington Monument. Further, we found that ____________ ADD FINDINGS ABOUT SIGNIFICANCE DURING SHOCK!!!

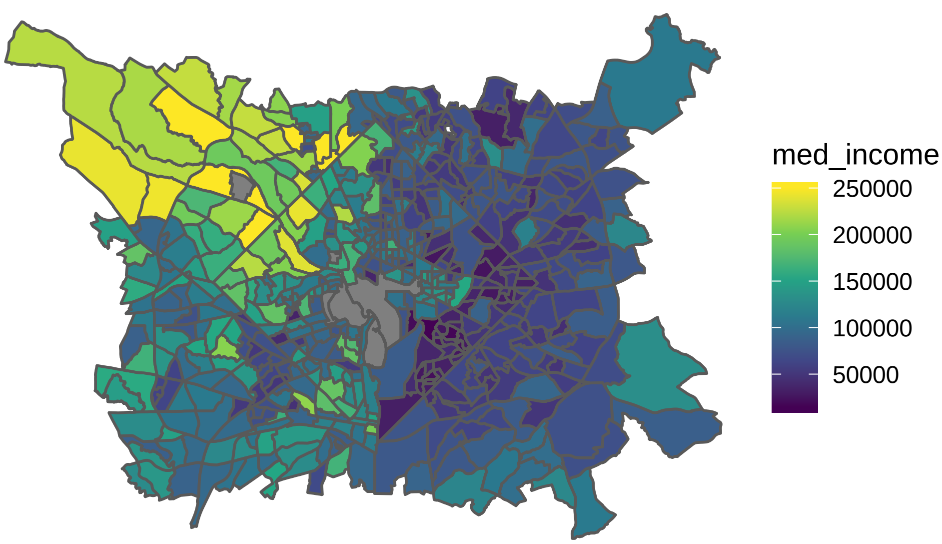
It is important to examine urban mobility in two ways. First, city attributes such as road or transit access make certain areas more accessible. Unfortunately, accessibility or mobility often falls along racial or class lines. Washington is highly segregated and by using a spatial autoregressive model on travel time to the Washington Monument during normal times, we are able to see the impact of race on D.C. area mobility. We observed that the population of black residents in a census tract is one of the few significant variables in predicting travel time to the Washington Monument. Second, major increases in population during special events can shock the city's infrastructure. For this reason, we looked at Washington because it is a city that is especially significant during major national events such as the inauguration and therefore attracts huge numbers of additional visitors. We used a similar spatial autoregressive model for the week of the inauguration and Women's March and then modeled the differences between the normal weeks and the inauguration and march to observe the shock on the system. Through modeling both before and during the major events, we were able to observe both demographic significance of travel time and the impact of major population changes on the system.

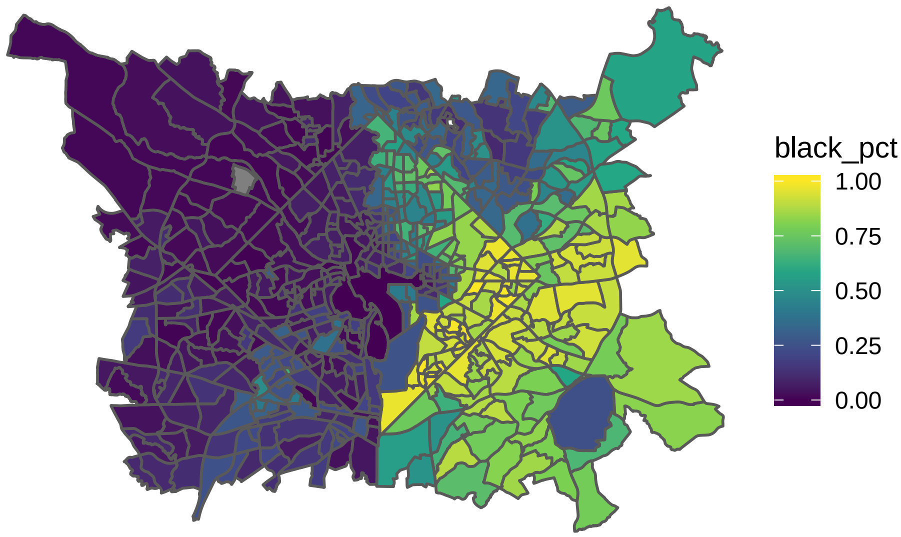
In order to examine the shock from these massive events, we used areal Uber data which includes travel times between various census tract locations in the DC metropolitan area. To model travel times before and during the event periods, we isolated a location and looked only at the travel time between each census tract in the Washington metropolitan area and the Washington Monument. We chose the Washington Monument as the origin because it is in the middle of the National Mall, where many significant events, including the inauguration and Women's March, took place. Because the shape files are census tracts, we were able to include census information for each polygon through the `tidycensus` package. This data allowed us to address the impact of demographics on travel time.

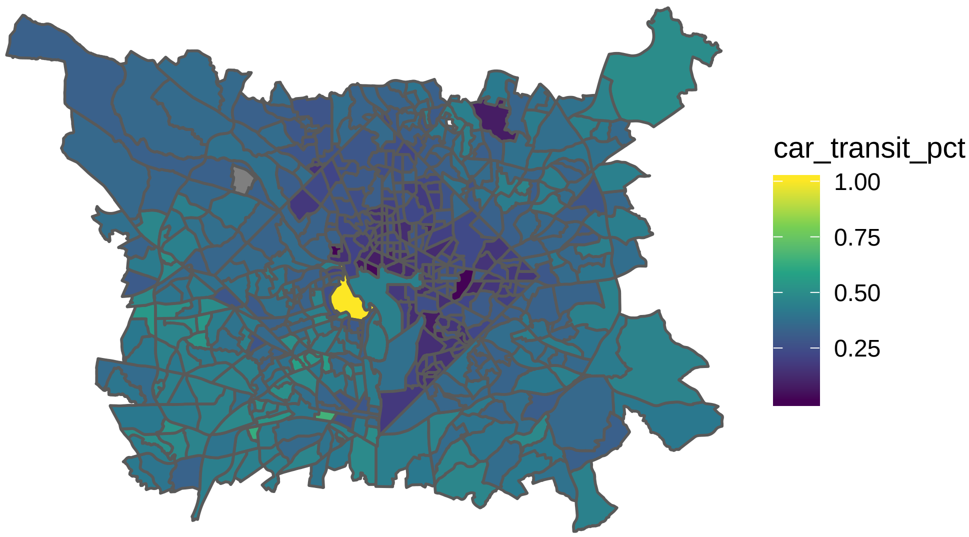
Literature Review

Many scholars have studied both mobility and shocks to the system in a variety of contexts. It is important to closely study every city's transportation system which is why many studies have been conducted. Noulas et. al use a network called Foursquare which records peoples locations to track and model mobility in a number of different cities (**???**). Studies about mobility and accessibility have been conducted in many cities including New York City (**???**) where concepts such as accessibility can be quantified and modeled to demonstrate lapses in the system. Shocks to transportation systems have also been modeled by a number of scholars. Donovan and Work's 2017 study uses taxi data to model the impacts of major events such as hurricane Sandy on New York city (**???**). Wang also modeled the impact of Hurricane Sandy on New York city by looking at images of people throughout the city (**???**). Other major events modeled by scholars include the Olympics (**???**) and the World Cup (**???**). Scholars have used a variety of data sources to observe the mobility of people including the taxi data and images as mentioned above, and cell phone data (**???**). Uber data is accessible and also provides an indication of mobility since so many people travel in cars. Because of this accessibility, we decided to model the inauguration and Women's March in Washington, D.C., two major events that occurred back to back.









Modeling

Normal Mean Time Model

```
##
## Call: spautolm(formula = MeanTimeNonTreat ~ med_income + asian_pct +
##     black_pct + white_pct + public_transit_pct + car_transit_pct +
##     n_roads + n_transit + distance, data = travel_time_diff,
##     listw = listW, family = "SAR")
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -380.99041  -67.50714   -0.10946   65.36708  426.82127
##
## Coefficients:
##                        Estimate  Std. Error z value  Pr(>|z|)
## (Intercept)          1.1579e+03  1.8395e+02  6.2946 3.081e-10
## med_income          -1.8695e-04  1.7990e-04 -1.0391  0.298739
## asian_pct           -2.0781e+02  1.3153e+02 -1.5800  0.114098
```

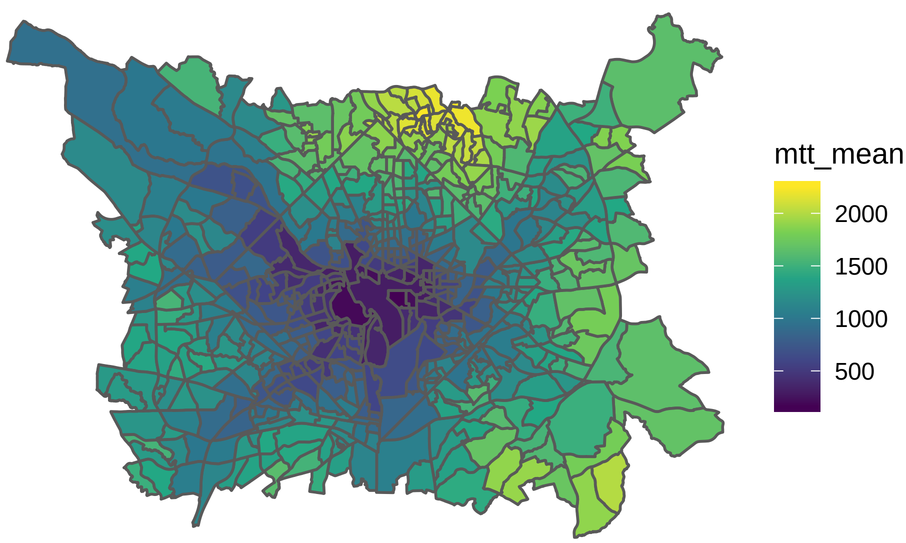
```
## black_pct          -1.0516e+01  7.5098e+01 -0.1400  0.888639
## white_pct          -3.1885e+01  7.8941e+01 -0.4039  0.686279
## public_transit_pct -3.4827e+01  9.9065e+01 -0.3516  0.725173
## car_transit_pct    -1.1675e+02  7.7121e+01 -1.5138  0.130064
## n_roads            -7.5302e+00  1.2009e+00 -6.2703 3.603e-10
## n_transit          -2.1267e+01  7.8448e+00 -2.7110  0.006709
## distance            2.0080e-02  6.7753e-03  2.9637  0.003039
##
## Lambda: 0.97056 LR test value: 911.85 p-value: < 2.22e-16
## Numerical Hessian standard error of lambda: 0.0099435
##
## Log likelihood: -3438.657
## ML residual variance (sigma squared): 11345, (sigma: 106.51)
## Number of observations: 551
## Number of parameters estimated: 12
## AIC: 6901.3
```

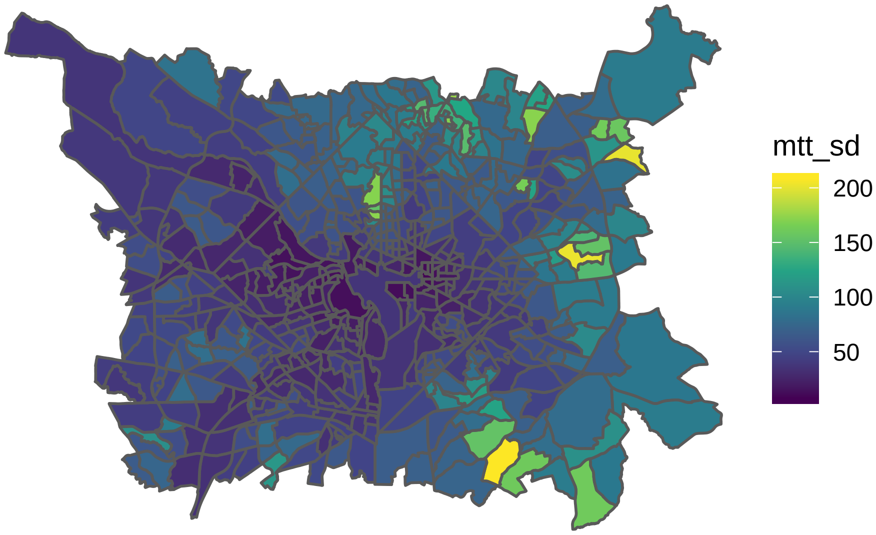
```
##
## Call: spautolm(formula = MeanTimeTreat ~ med_income + asian_pct + black_pct +
##     white_pct + public_transit_pct + car_transit_pct + n_roads +
##     n_transit + distance, data = travel_time_diff, listw = listW,
##     family = "SAR")
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -383.462207  -68.685301   -0.051871   65.139956  395.971092
##
## Coefficients:
##                      Estimate  Std. Error z value  Pr(>|z|)
## (Intercept)         1.1455e+03  1.6811e+02  6.8142 9.478e-12
## med_income         -2.2147e-04  1.8159e-04 -1.2196 0.2226108
## asian_pct          -2.2741e+02  1.3269e+02 -1.7138 0.0865679
## black_pct          -9.7893e+00  7.5733e+01 -0.1293 0.8971521
## white_pct          -2.9294e+01  7.9622e+01 -0.3679 0.7129398
## public_transit_pct -2.6048e+01  9.9940e+01 -0.2606 0.7943686
## car_transit_pct    -1.4030e+02  7.7824e+01 -1.8028 0.0714272
## n_roads            -7.0494e+00  1.2118e+00 -5.8171 5.986e-09
## n_transit          -2.1290e+01  7.9168e+00 -2.6892 0.0071619
## distance            2.2750e-02  6.7454e-03  3.3726 0.0007446
##
## Lambda: 0.96617 LR test value: 888.5 p-value: < 2.22e-16
## Numerical Hessian standard error of lambda: 0.01059
##
## Log likelihood: -3441.746
## ML residual variance (sigma squared): 11549, (sigma: 107.47)
## Number of observations: 551
## Number of parameters estimated: 12
## AIC: 6907.5
```

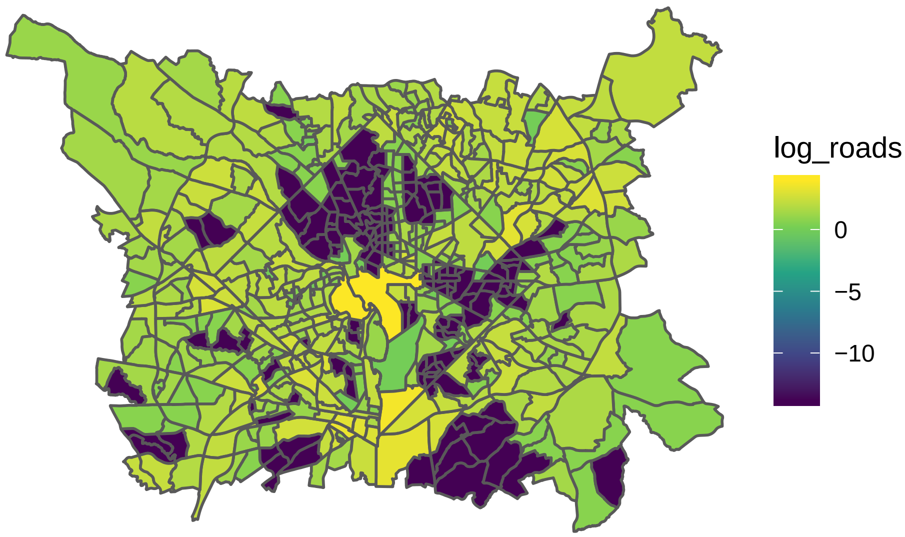
Difference Model

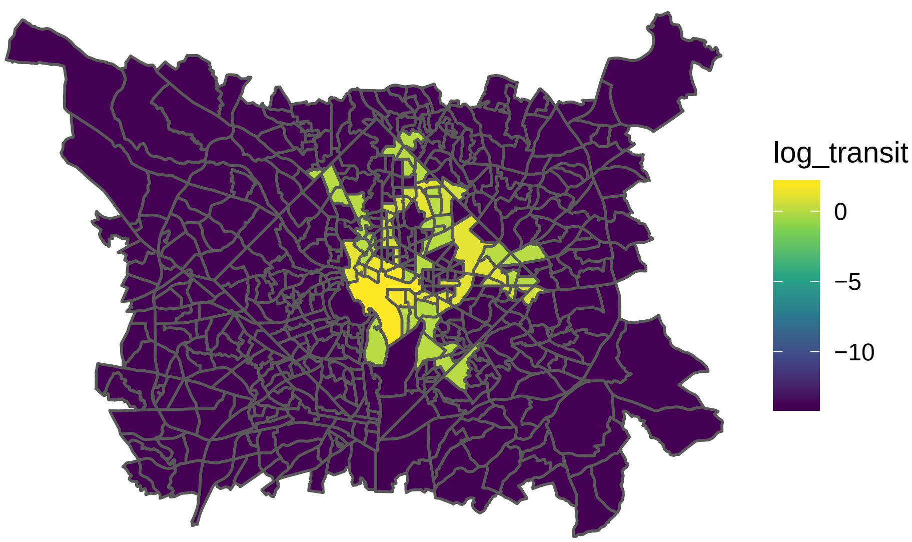
```
##
## Call: spautolm(formula = MeanTimeDiff ~ med_income + asian_pct + black_pct +
##     white_pct + public_transit_pct + car_transit_pct + n_roads +
##     n_transit + distance, data = travel_time_diff, listw = listW,
```

```
##     family = "SAR")
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -268.46503  -11.21467    0.41463   13.65852  235.26880
##
## Coefficients:
##                        Estimate  Std. Error z value Pr(>|z|)
## (Intercept)           4.5442e+01  2.6539e+01  1.7123  0.08685
## med_income           -2.7558e-05  6.7321e-05 -0.4094  0.68228
## asian_pct            -4.0027e+01  4.5094e+01 -0.8876  0.37474
## black_pct             3.8780e+00  2.4875e+01  0.1559  0.87611
## white_pct             1.1469e+01  2.7155e+01  0.4224  0.67276
## public_transit_pct  1.7912e+01  3.5314e+01  0.5072  0.61200
## car_transit_pct    -3.6233e+01  2.7889e+01 -1.2992  0.19387
## n_roads               6.5911e-01  4.3597e-01  1.5118  0.13058
## n_transit             2.8534e-01  2.8998e+00  0.0984  0.92162
## distance             -1.5541e-03  1.0361e-03 -1.5000  0.13361
##
## Lambda: 0.62662 LR test value: 122.91 p-value: < 2.22e-16
## Numerical Hessian standard error of lambda: 0.046651
##
## Log likelihood: -2838.64
## ML residual variance (sigma squared): 1606.8, (sigma: 40.085)
## Number of observations: 551
## Number of parameters estimated: 12
## AIC: 5701.3
```




mtt_sd
200
150
100
50


log_roads
0
−5
−10


log_transit
0
−5
−10

Results

Our first spatial autoregressive model looks at how mean travel times is affected by the estimated demographics in each census tract before the inauguration. The coefficients `med_income`, `asian_pct`, `public_transit_pct`, `car_transit_pct`, `n_roads`, `n_transit` are all negative, implying that it takes longer to travel from the Washington Monument to the polygons in which there are lower median income, lower percentage of Asian population, and lower number of public transit, car transit, roads, and metro lines. It is usually the case that the city attributes such as `public_transit_pct`, `car_transit_pct`, `n_roads`, and `n_transit` have an immediate impact on transportation and thus can directly account for the variation in travel times, whereas the link between demographic factors in a region and travel times might not be as clear. One possible explanation is that the areas where people have lower income are often further away from the city center and less accessible by public transport, which altogether results in longer travel times and decreased mobility. The coefficients `white_pct`, `black_pct`, and `distance` are all positive, implying that it is faster to get to the closer polygons or those in which there is a lower percentage of white and black residents from the Washington Monument. Although it makes intuitive sense that distance is positively correlated with time, it does not come as natural to think about why there might be such positive relationship between racial proportion of whites and blacks with travel times. Based on the visualizations, the fact that the predominantly black areas are seen further east while the predominantly white areas are seen further west indicates that the variation observed in mean travel times can be explained by the variation in geographical distance and inherent accessibility in areas dominated by certain races. Furthermore, it is not surprising that `n_roads`, `n_transit`, and `distance` all have significant p-values as they are generally the primary predictors that contribute to the estimation of travel times.

Now we consider our second spatial autoregressive model which looks at how mean travel times is affected by the estimated demographics in each census tract during and after inauguration. The coefficients from the second model stay consistent with our observation in the first model, that `white_pct`, `black_pct`, and `distance` are positive while `med_income_pct`, `asian_pct`,`public_transit_pct`, `car_transit_pct`, `n_roads`, `n_transit` are negative. Then, we fit a spatial autoregressive model of the difference in mean travel times between the two time periods with the demographic variables to see how urban mobility is affected by the influx of people coming to Washington D.C. during the shock events. Here, while most of the coefficients match with what we see above, `public_transit`, `n_roads`, and `n_transit` are indeed positive rather than negative. This implies that the difference in mean travel times increases as the number of roads in a polygon increases, which seems counterintuitive at first since we expect that these factors should increase transportation efficiency, however, rising traffic jam due to the surge in people coming to these massive events might be the potential cause of decreased mobility regardless. Overall, we find no indicator of a significant difference in travel times as well as a direct impact of the demographics on travel times before and after the shock events in Washington D.C.

Conclusions

Wallace, Tim, and Alicia Parlapiano. 2017. "Crowd Scientists Say Women's March in Washington Had 3 Times as Many People as Trump's Inauguration." *The New York Times*. The New York Times. https://www.nytimes.com/interactive/2017/01/22/us/politics/womens-march-trump-crowd-estimates.html?searchResultPosition=1.