Amazon\_fulfillment\_centers\_final.csv README:

This is the dataset I have compiled on the amazon fulfillment center network in the United States. It is not complete, but I believe that it is better than any dataset currently available on the internet. For reference this [nber paper](https://www.nber.org/papers/w23361.pdf) uses MWVPL.com for its dataset, while this [epi paper](https://www.epi.org/publication/unfulfilled-promises-amazon-warehouses-do-not-generate-broad-based-employment-growth/) uses a mwvpl.com, Avalara, and guidedimports.com and sports 54 observations. This dataset also made use of mwvpl.com.

Our dataset has 86 observations for both announced date of amazon center and opening date of that center, in addition to 148 opening dates, and 102 announced dates.

Errors:

This dataset does likely have some errors. While examining the MWVPL dataset, the one the nber/epi paper also used, I noticed that they would sometimes mistakenly attribute a news site announcing a planned center to the opening of that center. These errors should be resolved, as they are very noticeable when we have a date of announcement, and I went back and fixed the opening dates for those.

Any announcing/opening dates that seem too close, should be accurate, these are cases when Amazon moved into a fully operational warehouse.

Sources:

[Mwpvl.com](http://www.mwpvl.com/assets/images/autogen/Amazon-Fulfillment-Center-Network-in-USA.png), [Ecomtax](https://ecomtax.net/FBACenterMap), [taxjar blog](https://blog.taxjar.com/amazon-warehouse-locations/), [avalara](https://www.avalara.com/small-business/en/amazon-fulfillment-center-locations.html#NewJersey), [selleressentials](https://selleressentials.com/amazon/amazon-fulfillment-center-locations/), several other websites, and a lower bound of 1,000 newspaper articles.

What I did:

Transcribed this mwpvl.com image table to spreadsheet format, and then use other sources and newspaper articles to fill in the blanks. For example, I would search “Oregon announce amazon fulfillment center” filtered within a specific date range and then use address/ footage to match with data already in the dataset. (News articles often don’t have the code of the amazon center so for multiple in 1 city you have to piece it together using “first warehouse at 1000 road” “second at 2004 blvd”.)

Columns:

code – amazon fulfillment center unique identifier

sqft – square footage of the facility at announcement. Many facilities were later significantly expanded, but I have only recorded it for announcement.

announced - approximate date amazon center was announced in the news.

Opened - approximate date amazon center shipped its first package in the news. I have heavily relied on sources like mwvpl.com for many opening dates. Dates like 2016 Q3 or 2016 Fall have been converted to the middle month of that period, e.g. 11/1/2016.

Closing – when was the fulfillment center announced to be closed

Closed – when did the center close

accurate\_announced / accurate\_opened – Boolean. TRUE implies date is fairly accurate. (90% of data marked TRUE should be within 1 month). False implies date is correct to the year.