**4° Slide, RENT dataset:** The dataset has been built by scraping online rent announcements from the Craiglist website from 2011 to 2018. Most of the observations are not geolocated but we have information about the neighborhood they’re located in. “dire due parole sul fatto che ci sono buchi”

**5° Slide, EVICTION/BUYOUT dataset:** Evictions and Buyouts are considered to be proxies for gentrification, as often an increase in Evictions is observed in areas where the rent prices are rising. We’re interested in understanding how evictions and rent prices are related and if they lead to different conclusions.

The main information in the dataset are… . Addresses have been transformed to lat and long to be used in the models

**6° Slide, CONSTRUCTION dataset:** We’ll study how nearby constructions affect rent prices and Evictions amount. There are two possible interpretations we’ll investigate:

* Building new constructions increases the attractiveness of the nearby houses, driving up the prices
* Building new constructions increases the number of houses on the market, driving down the prices

**7° Slide, PARCEL dataset:** Based on previous papers on the topic, the effect of new constructions may be local and can’t be effectively detected by studying at a neighborhood granularity.

**8° Slide, MAPPA-VIDEO:** In this insightful animation we see the evolution of rent prices and eviction locations during 2011-2018. It is possible to see that the density of evictions is higher where rent prices tend to be higher.

**9° Slide, RENT functions:** From the picture here in the top we can see that there are weeks in which no rent announcements are collected. We performed a Gaussian Kernel Smoothing on rent observations and here we have the results. We also considered the first derivative, as we are particularly interested in the change of rent prices.

**10° Slide, OUTLIER DETECTION RENT:** Outlier detection is performed, looking for neighborhoods with an atypical behavior.

From the study on rent prices, there is one nhood that is behaving differently: Treasure Island. Treasure Island is both separated from the rest of the nhoods, being in fact an island, and have few observations. We decide to remove it from the study. There are other 2 nhoods that are borderline outliers: Western Addition and Lakeshore. We decide to keep them as they’re two central nhoods in SF geography, and we’ll eventually change our mind later.

From the study on the derivative of rent prices, again Western Addition is labeled as an outlier.

**11° Slide, EVICTIONS Functions**: We use the same approach used on Rent. It does not look like it behaves similarly to the rent smoothed functions. We’ll try to recompute them by normalizing on the area of each nhood.

**12° Slide, TEST ON RENT FUNCTIONS:** Here are reported 4 images describing 4 functional tests we performed. We clustered the nhood by simply computing the nhoods that have the highest number of constructions and evictions considering the whole timeframe. All p-values are high and so we conclude there is not a statistically significant difference between the distributions of the 2 groups.Proveremo a migliorare le nostre partizioni, ad esempio facendo clustering funzionale.

**13° Slide, GAM Model:** We start now analyzing the effect of new constructions on rent prices. Assuming there is a time lag between the permits of a new construction being emitted and an effect on rent prices, we included also new\_constructions up to 4 years before the rent online announcement.

The model has an R^2 of 0.36. Among the most significant variables there are “distance to financial district” and “year”. Covariates related to new\_constructions are hardly significant, denoting that at a nhood granularity they’re not significantly affecting the rent prices.

**MENTRE vado avanti con le slides:** We’ll move to a parcel granularity. In this way we’ll be able to detect local effect (up to 100m) of new\_constructions on rent prices.