1. Binowanie biznesów i restauracji do rastrów
2. Doklejenie info (odległość od centrum, gęstość przystanków, gęstość dróg)
   1. Pobranie Warszawy z API (<https://github.com/ropensci/osmdata>)
   2. Zmiana projekcji
   3. Binowanie liczby przystanków
   4. Obliczenie długości dróg w kwadraciku (https://gis.stackexchange.com/questions/280760/intersecting-lines-and-polygons-and-calculating-line-length-in-r)
3. EDA
   1. sprawdzenie korelacji pomiędzy zmiennymi bez modelu
   2. korelacja przestrzenna pomiędzy wszystkimi zmiennymi (Moran’s I powinno działać chyba)
4. Do każdego kwadracika doklejenie zmiennych zawierających info o sąsiednich
   1. Dla każdego kwadracika znajdź id sąsiadów
   2. Do kwadracika dodaj info o tym czy jest skrajny i ile ma sąsiadów
   3. Pobierz dane o sąsiadach danego kwadracika do data frama
   4. Agreguj dane i zapisz do zmiennych w kwadraciku
5. **Klasyfikacja kwadracików czy zawierają restauracje za pomocą zmiennych wśród których jest liczba mieszkańców i ilość biznesów**
6. Model ostateczny ze wszystkim dodanym:
   1. **MSAR- jest zarówno zmienna zależna jak i zmienne niezależne pobrane od sąsiadów**
   2. Zmienna zależna- czy w danym rastrze jest >1 restauracja (0-1)
   3. Info na raster:
      1. Odległość od centrum
      2. Gęstość dróg
      3. Ilość przystanków
   4. Info na raster- badane:
      1. Ilość biznesów w rastrze
      2. Ilość mieszkańców w rastrze
   5. Info spatial
      1. Wszystkie w obrębie 8 kratek dookoła
      2. Zmienna zależna- suma z kratek
      3. Ilość biznesów i mieszkańców w rastrze+gęstość dróg i przystanków- suma, min, max z kratek
7. Zbadać co ma większy wpływ- biznes czy ilość mieszkańców czy nic
   1. Wyestymowanie null modelu bez zmiennych z hipotezy
   2. Wyestymowanie modelu ze wszystkimi zmiennymi i Varimp – najprostsze
   3. Estymacja modeli z i bez danej zmiennej i porównanie wyników
8. Użycie 2 modeli- spatial logistic regression i random forest
   1. Praca o logistic [h*ttps://pdfs.semanticscholar.org/8ed8/c0497997e5831ecdfc01d8df93b7b6216c98.pdf*](https://pdfs.semanticscholar.org/8ed8/c0497997e5831ecdfc01d8df93b7b6216c98.pdf)
   2. GWLR- GWmodel::ggwr.basic
   3. Random forest normalnie jak w caret
9. *cross-validation za pomocą blockCV czy zwykłe losowanie?- kiedy są już dodane do bazy zmienne określające sąsiedztwo*

Tekst alternatywny wygenerowany przez komputer:
The general form of the model considered here is as follows (eq.l): 
Y— pWY+atN+ Xß+ wxe+ u and u—AWu+ E 
(1) 
where Y is the vector of individual rental transactions (dependent variable), pWY is the 
average rental fee in a neighbourhood defined with the spatial weights matrix W (spatial lag 
of the dependent variable), is a constant term, Xß is the set of explanatory variables for 
a given transaction, WXO is the set of spatial lags (average values weighted with W in a 