chiarezza obiettivi: tra il nostro 3o e quello che diceva lui, chiedergli cosa intendesse fare.

domande sul dataset:

- differenza tra raw\_stop\_counts e raw\_device\_counts:

Assunzione 1 device associata a un individuo

- stops\_by\_day vs stops\_by\_hour

- device\_daytime\_areas e work\_hours\_device\_home\_areas

- cosa è uno stop? se mi muovo ma rimango nello stesso CBG questo conta come un nuovo stop?

Metadati: carichi integrativi come insediazione di gruppi

Un potenziale generativo su come sono generate le census block

Verificare se (avendo i codici di) raggruppare i census per territori/città. Lasciare le agglomerazioni urbane e togliere quelle periferiche.

Obiettivo:

Sulla base dei dati prevedere dove si trova la permanenza

Clusterizzazione di spazi per la natura di movimenti. Continuità spaziale può sovrapporsi ai dati? Si pensa di si

Mondo legato alla network analisi. Potenziale di attrazione, densità di popolazione che si può estrarre dal census. Carico insediativo può essere considerato per attrarre

20 punti di interessi ma

1- predizione sul tempo stimato in un luogo —> densità di popolazione

2- vedere se le i census block piú

3- consulenza negozi (anche solo pubblicitaria) —>

**post call**

* regressione lineare, lasso cross validation cose sul primo obiettivo facile (+ chiedere la densità di popolazione dei metadata)
* è vero che i blocchi vicini con time stops simili formano una città? (non punto principale, secondo momento)
* dopo ridge regression: parliamo di classificazione e clustering -> sulle marche? / bisogna capire come sfruttare quei dati. suddividere i brand legati ad ambiti diversi. dividere i brand che abbiamo in classi, capire come interpretare qst dato. sai che in quella zona lì c’è o meno bisogno di una classe di brand(o di un brand specifico)
* su qst, occhio alla vicinanza tra i blocchi

Cose da fare:

1. creazione di mappa in R (nelle vari {CBG: n. devices} std::map<CBG, n. devices> MAPPA

MAPPA[key] -> value

così come per i vettori, per accedere agli elementi fai vettore[index]

-> cercare un tipo di var

**fatto:** abbiamo creato un vettore coi n devices dove ogni elemento ha come nome il CBG corrispondente. vd questo sito per capire come funziona l’accesso alle variabili

http://www.r-tutor.com/r-introduction/vector/named-vector-members

1. cercare qualcosa in letteratura
2. Capire di quali metadati abbiamo bisogno (leggere census\_metadata)

**COSE DA FARE DA QUI A 22/4**

**da fare**

**1. the general problem:** (frasi motivazionali sito e anche qua in relazione al NOSTRO OBIETTIVO)

**2. the data :** (descrizione del dataset, in base a cosa abbiamo trovato di interessante o cosa vorremmo analizzare)

**3. the research objectives and expected results :**

analisi predittiva e roba negozi(da capire bene)

**4. analysis tools used (among those treated so far in the course) and the results obtained**

vd sotto

**5. any open questions to be possibly addressed to those present**

* covid: sentire Izzo , scoprire date e fasi americane cosa sono
* regressione lineare -> gaussianità delle y?? che y???
* cluster: se dei brand come però? non è facile compararle per ora,che distanza usiamo
* cluster: raggruppi i cbg come densità popolazione, periferia o no
* domande sito:

**descriptive**

‒ which day of the week is a CBG busiest?

one way anova (normalizzando rispetto alla densità di popolazione… magari si possono confrontare tutte le CBG)

oppure selezioniamo solo quelle delle città e fai analisi lì (e poi campagna)

-> chiedere a Izzo qst dato campagna/città **dice di cercare noi un threshold nelle densità che ci calcoliamo noi ...**

‒ when during the day is a CBG busy?

‒ where do the devices that stop in a CBG during breakfast, lunch or dinner time travel from

**predictive**(support to location intelligence): predict the time of stay (median\_dwell) for a suitable allocation of services / businesses to catch, pedestrian flows slide …

regressione lineare qua