**Notes for Meeting**

**Time:** 2021.03.11 7:00-8:00 PM

**Participants:**

Min Lin: blue Brisilda Ndreka: yellow Chrstine: red

**Main Topics:**

**1.Presentation by Min:**

Tidy data (see R markdown about the definition of a tidy data, 3 basic elements)

Main point: Make data transformation and exchange the columns

**2.Check on the datasets I (Excel forms)**

Reach an agreement: Save data from Excel to R without losing any information, include a new column Report\_Percent <this number is an approximation average of data from Area actually reporting divided Estimated Total for each crime type>

(Estimated Total rows being deleted)

**3.Check on the datasets II (Excel forms)**

Check the Crime Type columns:

Violent Crime=Murder+ Rape(revised) +Robbery+ Aggravated assault

Delete “Violent Crime”

Keep both columns of Rape to check two definitions of Rape

Property Crime=Burglary+ Larceny-theft + Motor

Delete “Property Crime”

**4.Links mentioned in the meeting:**

R-gallery

<https://www.r-graph-gallery.com/>

(Time series menu)

**\*\*\*5. Clean the data in Excel <no better idea rather than hand cleaning so far>**

Cleaning the headings (3 rows)

Includes the following in your new dataset for every year:

State, Area, repor\_percent, population, Four Violent Crime, Three Property Crime

Notes: Add V\_ before each violent crime

Add P\_ before each property crime

**6.Further work ideas:**

Create statistical models to predict crime?

Ideas: Multinomial model--Response as number of crime type, population, area, GDP and other information

Construct an Evaluation System: PCA method

**To do next and assign the text:**

**High Priority:**

Clean the data for each year in Excel a tidy form (1999~2005)(2006~2012)(2013~2019)

**Medium Priority:**

Information for 1995~1998 in .txt to ask Haim Bar

Put all the data in a R package (further work)