**Assignment 1 – ECON613 Spring 2019**

David Christian

**Exercise 1 – Missing Data**

Report the following statistics

* Number of students: 340,823 students
* Number of schools: 641 schools
* Number of programs: 33 programs
* Number of choices (school,program): 3,086 choices (school,program)
* Missing test score: 179,887 students
* Apply to the same school (different programs):
  + Interpretation 1: number of students who apply to at least two different programs in at least one same school:
* Apply to less than 6 choices: 17,088 students

**Exercise 2 – Data**

Create a school level dataset, where each now corresponds to a (school.program) with the following variables:

* The district where the school is located
* The latitude of the district
* The longitude of the district
* Cutoff (the lowest score to be admitted)
* Quality (the average score of the students admitted)
* Size (number of students admitted)

It is in the R command.

**Exercise 3 – Distance**

Using the formula, calculate the distance between junior high school and senior high school (of each students and each of their choices)!

**Exercise 4 – Descriptive Characteristics**

Report the average and sd of the cutoff and quality for each ranked choice and later by student test score quintiles

|  |  |  |  |
| --- | --- | --- | --- |
| By (rank) | Cutoff | Quality avg | Quality sd |
| 1 | 165 | 314 | 56.4 |
| 2 | 173 | 302 | 49.0 |
| 3 | 190 | 289 | 42.4 |
| 4 | 185 | 277 | 37.5 |
| 5 | 198 | 253 | 30.4 |
| 6 | 158 | 251 | 28.9 |

|  |  |  |  |
| --- | --- | --- | --- |
| By (score quintiles) | Cutoff (min admitted score) | Quality avg | Quality sd |
| 1st (lowest) | 158 | 214 | 5.69 |
| 2nd | 221 | 254 | 16.9 |
| 3rd | 283 | 310 | 17.6 |
| 4th | 345 | 369 | 17.0 |
| 5th (highest) | 407 | 422 | 12.3 |

**Exercise 5 – Diversification**

Group schools by decile of selectivity (cutoffs), and compute for each individual the number of groups in the application. Redo this, by student test score (quintile).