DS1EDP: Homework 03 – Solutions

1. Differences between Universitites

Question 1:

```
dissimilarity = sum(abs(stanford – berkeley))
```

Question 2:

```
revised_dissimilarity = sum(abs(stanford - berkeley) * weights)
```

2. Unemployment

Question 1:

```
unemployment = Table.read_table("unemployment.csv")
```

Question 2:

```
by_nei = unemployment.sort("NEI", descending=True)
by_nei_pter = unemployment.sort("NEI-PTER", descending=True)
```

Question 3:

```
greatest_nei = by_nei.take(np.arange(10))
```

Question 4:

```
pter = unemployment.column("NEI-PTER") – unemployment.column("NEI")
```

Question 5:

```
by_pter = unemployment.with_column("PTER", pter).sort("PTER",
descending=True)
```

Question 6:

```
year = 1994 + np.arange(by_pter.num_rows)/4
pter_over_time = unemployment.with_column("Year", year, "PTER", pter)
pter_over_time.plot("Year", "PTER")
```

Question 7:

```
highPTER = True
```

3.   Birth Rates

Question 1:

```
us_birth_rate = sum(pop.column("BIRTHS")) / sum(pop.column("2015"))
```

Question 2:

```
growth_rates = pop.with_column('Growth Rate', (pop.column(3) /
pop.column(2)) - 1)
fastest_growth = growth_rates.sort("Growth Rate",
descending=True).take(np.arange(5)).column("NAME")
```

Question 3:

```
migration_rates = pop.with_column("ARM",
abs(pop.column("MIGRATION") / pop.column("2015")))
movers = migration_rates.where("ARM", are.above(0.01)).num_rows
```

Question 4:

```
west_births = sum(pop.where("REGION",
are.equal_to("4")).column("BIRTHS"))
```

Question 5:

```
less_than_west_births = pop.where("2016",
are.below(west_births)).num_rows
```

Question 6:

```
correlation = True
```

4.   Marginal Histograms

Question 1:

```
option = 2
```