

ECON 21020 PSET 3

Allen Zhou

10/29/2023

(a) 420 Observations

(b)

- i. The variable *income* measures the average district income in dollars
- ii. Mean = 15.31659, STD = 7.22589
- iii. Mean = 15316.59, STD = 7225.89. These values are expected as $income = avginc * 1000$ so the mean and std should also be 1000 times larger as well

```
ds <- mutate(ds, income = avginc * 1000)

mean_avginc <- mean(ds$avginc)
std_avginc <- sd(ds$avginc)
print(mean_avginc)
```

```
## [1] 15.31659
```

```
print(std_avginc)
```

```
## [1] 7.22589
```

```
mean_income <- mean(ds$income)
std_income <- sd(ds$income)
print(mean_income)
```

```
## [1] 15316.59
```

```
print(std_income)
```

```
## [1] 7225.89
```

(c)

- i. The mean math score across all districts is 653.3426
- ii. Fraction of districts = 0.5785714, mean math score = 655.7177
- iii. Fraction of districts = 0.4214286, mean math score = 650.0819
- iv. Multiplying the fraction with the mean math score of each respective districts yields in the mean math score across all districts

- v. Let M_s be the mean math score for class size 20 or fewer and let M_l be the mean math score for class size larger than 20. $H_0: M_s = M_l$ or $M_s - M_l = 0$ $H_1: M_s \neq M_l$ t-statistic: 3.121799 p-value: 0.001922331 Since the p-value is less than $\alpha = 0.1$, and the t-statistic is greater than the absolute value of the critical value, we reject the null hypothesis
- vi. Covariance between *avginc* and mean math score is 94.7795. Covariance between *income* and mean math score is 94779.5. The values are not the same, but differ by a multiple of 1000, which makes sense since *income* is *avginc* * 1000
- vii. Correlation between *avginc* and mean math score is 0.6993981. Correlation between *income* and mean math score is also 0.6993981. The two correlations are the same because the two are measuring the same relationship between average income and mean math score, the only difference being the units, which doesn't come into play when measuring correlation.

```
## [1] 653.3426

## [1] 0.5785714

## [1] 655.7177

## [1] 0.4214286

## [1] 650.0819

## [1] 653.3426

## [1] 3.121799

## [1] 0.001922331

## [1] -1.648507

##
## Welch Two Sample t-test
##
## data: class_size$math_scr and class_large$math_scr
## t = 3.1218, df = 399.45, p-value = 0.001928
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 2.086702 9.184839
## sample estimates:
## mean of x mean of y
## 655.7177 650.0819

## [1] 94.7795

## [1] 94779.5

## [1] 0.6993981

## [1] 0.6993981

““
```