'DO THE GALAXIES FARTHEST TO THE EARTH DIFFER SIGNIFICANTLY FROM THE CLOSEST ONES IN TERMS OF THEIR TOTAL LUMINOSITY?'

answered, using hypothesis test

by Riyad Valiyev
TUT0101 Group #6

Samples of the furthest and closest galaxies

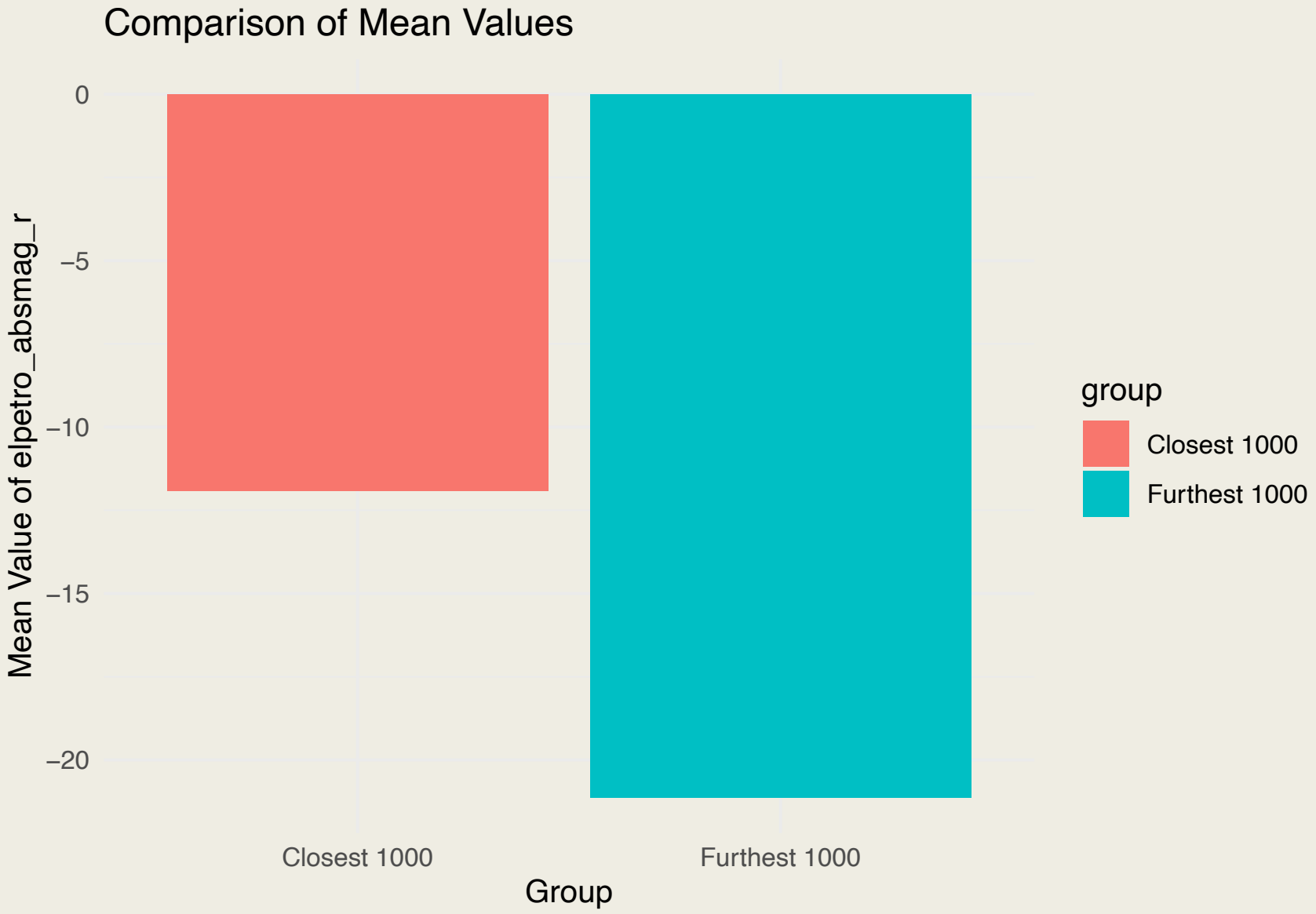
```
## Rows: 1,000
## Columns: 3
## $ elpetro_absmag_r <dbl> -21.93583, -21.48143, -19.69226, -21.43395, -22.41922~
## $ redshift         <dbl> 0.1499996, 0.1499991, 0.1499987, 0.1499983, 0.1499975~
## $ distance         <chr> "furthest", "furthest", "furthest", "furthest", "furt~
```

The ***furthest*** 1000 galaxies with their associated redshift and total brightness values

```
## Rows: 1,000
## Columns: 3
## $ elpetro_absmag_r <dbl> -0.8244038, -1.1099548, -6.2743492, -12.0312986, -9.2~
## $ redshift         <dbl> -0.0044815368, -0.0035639217, -0.0033348468, -0.00312~
## $ distance         <chr> "closest", "closest", "closest", "closest", "closest"~
```

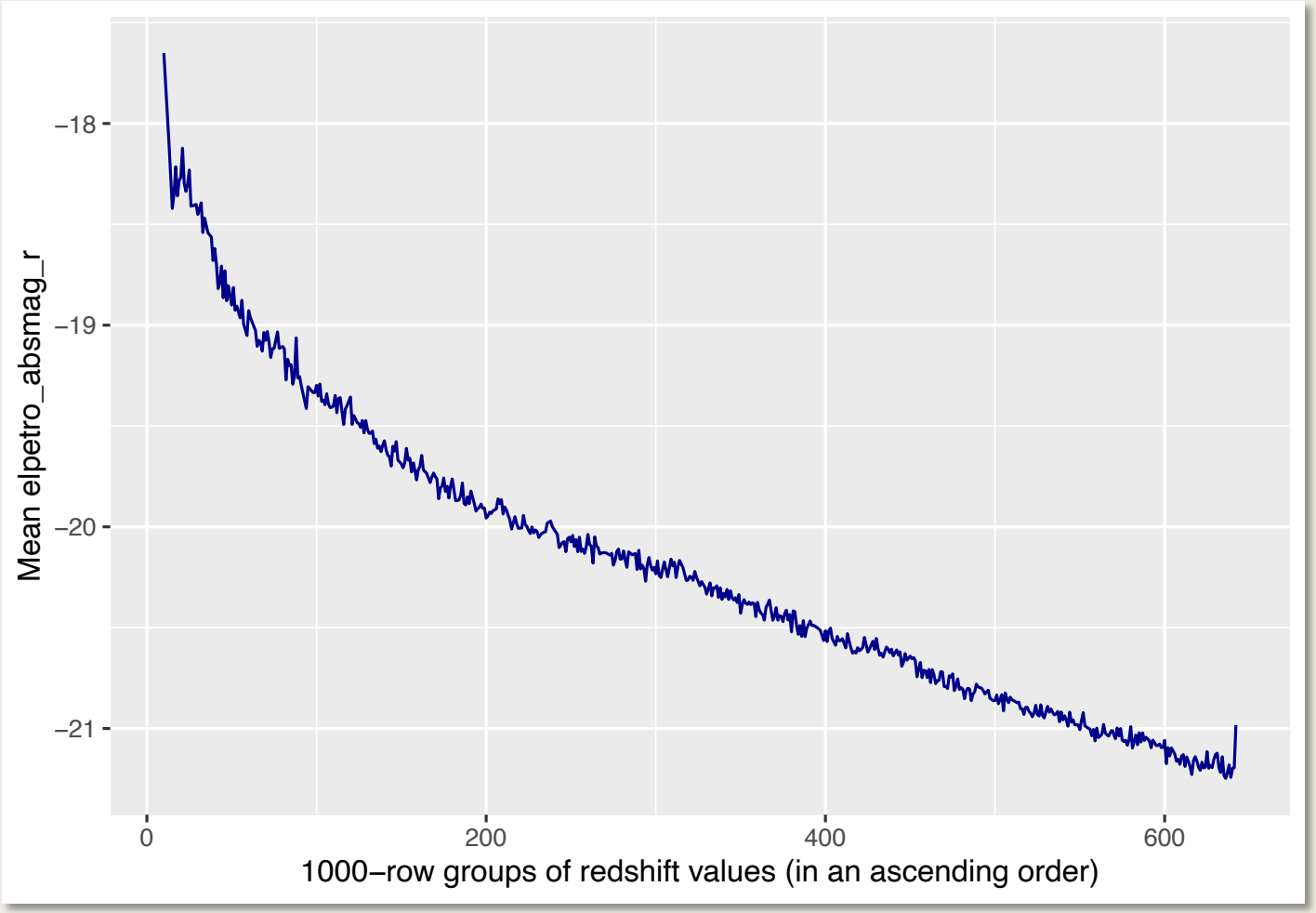
The ***closest*** 1000 galaxies with their associated redshift and total brightness values

Measured mean total brightness of each group



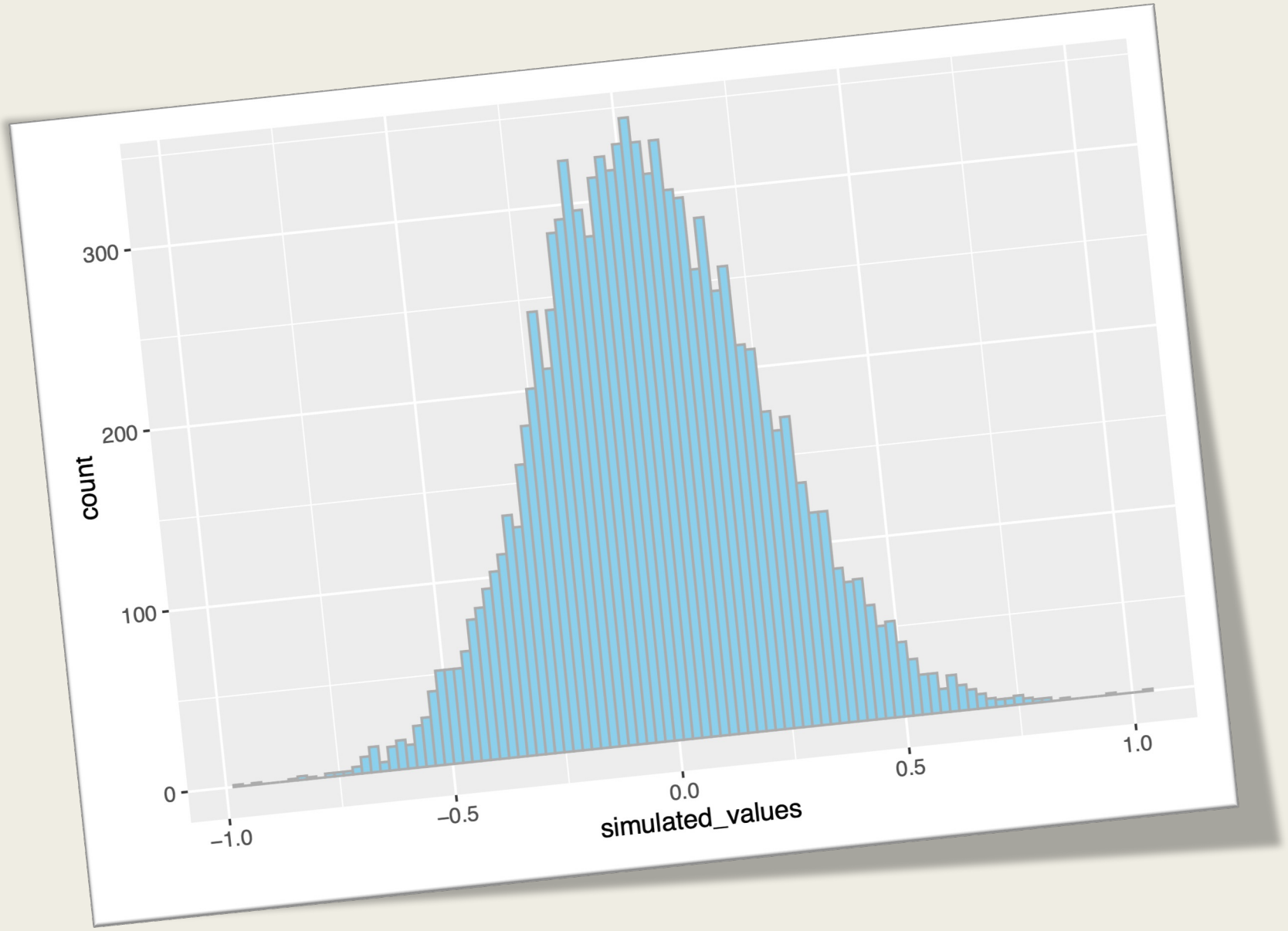
Correlation between redshift and total brightness

- Arranged 641,166 galaxies in an ascending order of their redshift values.
- Clustered them in the groups of 1000.
- Computed the mean “total brightness” of each group.
- Removed the outliers (only kept the groups with means in IQR).
- Plotted the result with a line graph.



Galaxy groups with higher redshift have less total brightness.

Hypothesis Test Results: Zero p-value!



```r
# compute the p-value
  # If you are performing a 2-sided test we only care about the
  # *absolute difference* between the two samples, not whether the difference
  # is bigger or smaller
  p_2side = sum(abs(simulated_values) >= abs(test_stat)) / repetitions
  p_2side
```

```
## [1] 0
```