Answer Sheet for Assignment: Introduction to working with R/RStudio

Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

*Answer the questions below based on your script output*

|  |  |  |
| --- | --- | --- |
| Question | | Answer |
| 1 | How many total flights (including cancelled flights) have Philadelphia (PHL) as the destination airport during December 2017? |  |
| 2 | What was the average arrival delay (in minutes) across all flights during December 2017? |  |
| 3 | What was the average arrival delay (in minutes) for American Airlines (with UniqueCarrier code of AA) during December 2017? |  |
| 4 | What was the longest arrival delay for United Airlines (with UniqueCarrier code of UA) during December 2017? |  |
| 5 | On average, which airline (using UniqueCarrier) experienced greater arrival delays: American Airline (AA) or United Airlines (UA)? |  |
| 6 | For question #5, was this difference statistically significant? What is the p-value? (answer both questions in the blank to the right) |  |
| 7 | Which airline(s) had longest average flight distance? (you can list more than one if it’s a tie) |  |
| 8 | Which airline (s) had shortest average flight distance?  (you can list more than one if it’s a tie) |  |
| 9 | On average, which origin airport (using Origin) experienced greater taxi out times: Newark (EWR) or Philadelphia (PHL)? |  |
| 10 | For question #9, was this difference statistically significant? What is the p-value? (answer both questions in the blank to the right) |  |
| 11 | Looking at the histogram. Is the distribution symmetric? Are most flights delayed less than 50 minutes or more than 50 minutes? |  |