**Documented changes from R to Python**

|  |  |  |  |
| --- | --- | --- | --- |
| **R** | **Python** | **Purpose** | **Differences** |
| filter() | drop() | Remove observations by values or conditions | Equivalent |
| arrange() | sort\_values() | Reorder rows by column or conditions | Includes sort\_values(by=’column’, ascending=True); True is default |
| select() | loc[] | Pick variables by column or row names/conditions | Some different functionality; do more string operations in Python |
| mutate() | assign() | Create new variables with functions of existing variables | Pretty similar |
| summarize() | describe() | Collapse many values into a single summary | Equivalent |
| group\_by() | groupby() | Group values by some variable | Not large difference; some chaining in Python must be done before or after, check documentation when necessary |
| str(data) | info() | Provides information on data and variables | Equivalent |

**Changes: Class Notes**

* replaced mpg dataset with some other dataset (TBD)
* used string operations to select columns that start with ‘d’ rather than select()

**Changes: Lab**

* demonstrated splicing
* added comments explaining operational differences between R and Python depending on what had taken place