

# Exercise

April 20, 2022

**0.0.1 Step 1. Import the necessary libraries**

**0.0.2 Step 2. Read dataset, and assign it to a variable called users and use the 'user\_id' as index**

```
[2]: ... = 'https://raw.githubusercontent.com/justmarkham/DAT8/master/data/u.user'
users = ...(..., sep='|', index_col='user_id')
```

**0.0.3 Step 3. See the first 25 entries**

```
[ ]:
```

**0.0.4 Step 4. See the last 10 entries**

```
[ ]:
```

**0.0.5 Step 5. What is the number of observations in the dataset?**

```
[ ]:
```

**0.0.6 Step 6. What is the number of columns in the dataset?**

```
[ ]:
```

**0.0.7 Step 7. Print the name of all the columns.**

```
[ ]:
```

**0.0.8 Step 8. How is the dataset indexed?**

[ ]:

**0.0.9 Step 9. What is the data type of each column?**

[ ]:

**0.0.10 Step 10. Print only the occupation column**

[ ]:

**0.0.11 Step 11. How many different occupations are in this dataset?**

[ ]:

**0.0.12 Step 12. What is the most frequent occupation?**

[ ]:

**0.0.13 Step 13. Summarize the DataFrame.**

[ ]:

**0.0.14 Step 14. Summarize all the columns**

[ ]:

**0.0.15 Step 15. Summarize only the occupation column**

[ ]:

**0.0.16 Step 16. What is the mean age of users?**

[ ]:

**0.0.17 Step 17. What is the age with least occurrence?**

[ ]: