Tabular Data and Pandas and Scipy

Biomedical Data Analysis in Python3

All materials and slides are available on GitHub (ZaneMuir/FDU-DataAnalysis-Workshop).

- Introduction
- More on File Types of Tabular Data
- ***Designing a Table***
- ***Querying and Grouping***
- Pandas demo

Introduction

Name	Gender	Age	Occupation
Morty Smith	Male	14	Student
Summer Smith	Female	16	Student
Beth Sanchez	Female	38	Horse Surgeon
Jerry Smith	Male	38	n/a

1	3	4	2
9	2	1	2
3	NaN	2	3
12	4	22	4
7	6	4	2

- **2-dimension**
- **More than Numbers**
- **A Header (Names)**
- **More than One Physical Unit**

- **Multi-dimension**
- **Only Numbers**
- **No Header (Names)**
- **Same Physical Unit**

Tabular Data File Formats

- Never use **Excel** for your python data analysis
  (or any other programming language)
- **CSV - Comma Separated Values**
- SQL - Structured Query Language (Sqlite, MySQL, etc.)
- NoSQL (MongoDB, etc)

Name	Gender	Age	Occupation
Morty	Male	14	Student
Summer	Female	16	Student
Beth	Female	35	Horse Surgeon
Jerry	Male	n/a	Ad Agent

```
Name,Gender,Age,Occupation
Morty,Male,14,Student
Summer,Female,16,Student
Beth,Female,35,Horse Surgeon
Jerry,Male,,Ad Agent
```

Designing a Table

- **Never Use Column Header**
- **Row Header as Abstract Concepts Only**

Name	Gender	Age	Occupation
Morty Smith	Male	14	Student
Summer Smith	Female	16	Student
Beth Sanchez	Female	38	Horse Surgeon
Jerry Smith	Male	38	Ad. Agent

Designing a Table Example

- **Weights of mice**

- **Behavior tasks**

Mice Weights

Group	Mouse	2019-01-03	2019-01-04	…
Exp	PV-Cre-034	23.4	23.4	…
	SOM-Cre-066	17.4	17.4	…
	…	…	…	…
Shm	SOM-Cre-064	18.9	18.9	…
	PV-Cre-035	22.3	22.3	…
	…	…	…	…

Mouse	Date	Weight	Group
PV-Cre-034	2019-01-03	23.4	Shm
PV-Cre-035	2019-01-03	22.3	Exp
SOM-Cre-064	2019-01-03	18.9	Shm
SOM-Cre-066	2019-01-03	17.4	Exp
PV-Cre-034	2019-01-04	23.4	Shm
PV-Cre-035	2019-01-04	22.3	Exp
SOM-Cre-064	2019-01-04	18.9	Shm
SOM-Cre-066	2019-01-04	17.4	Exp
…	…	…	…

Behavior Tasks

Day 1

Trial	1	2	3	4	5	6	7	8	9	10
Mouse1	O	X	O	X	X	O	O	X	X	X
Mouse2	O	X	O	X	O	X	O	X	X	X
Mouse3	O	O	X	O	X	O	X	X	O	X

Day 2

Trial	1	2	3	4	5	6	7	8	9	10
Mouse1	O	X	O	X	X	O	O	X	X	X
Mouse2	O	X	O	X	O	X	O	X	X	X
Mouse3	O	O	X	O	X	O	X	X	O	X

... ...

Mouse	Date	Trial	Performance
1	1	1	X
1	1	2	O
1	1	3	X
1	1	4	X
...	...	...	...
3	2	16	X
3	2	17	O
3	2	18	X
3	2	19	X
3	2	20	O

Pandas — Demo

Querying and Grouping