

Omics Central

Amrit Singh

2020-01-08

Contents

1	Rationale	5
2	Introduction	7
3	Data-types	9
3.1	Microarrays	9
3.2	RNA sequencing	9
3.3	Nanostring	9
3.4	Biocrates	9
3.5	Multiple Reaction Monitoring	9
4	Exploratory Data Analysis	11
4.1	Principal Component Analysis	12
4.2	K-Means	12
4.3	Hierarchical clustering	12
4.4	Sample plots	12
4.5	Variable plots	12
5	Batch Correction	13
5.1	ComBat	13
5.2	Surrogate Variable Analysis	13
5.3	Model adjustment	13
6	Differential Expression Analysis	15
6.1	Ordinary Least Squares	15
6.2	Linear Models for MicroArrays and RNA-Seq	15
6.3	Significance Analysis for Microarrays (SAM)	15
6.4	cell-specific Analysis for Microarrays (csSAM)	15
7	Network Analysis	17
7.1	DINGO	17
7.2	WGCNA	17
7.3	PANDA	17
7.4	BioNetStat	17

8	Data Integration	19
8.1	Supervised	19
8.2	Unsupervised	19
9	Biological Enrichment	21
9.1	Enrichr	21
9.2	SEAR	21
9.3	CAMERA	21
9.4	Network-based Gene Set Analysis	21
10	Literature Mining	23

Chapter 1

Rationale

This project was developed in order to create a resource warehouse for researchers analyze omics datasets of various types such as transcriptomics, proteomics, metabolomics. I expect this resource to grow as others contribute to it. Think of it as an awesome-resource github repo but in a bookdown format. However, since this book is meant as documentation to the omics central web application, adding new methods will require pull requests to the omics central web app repos (omics-central-frontend, omics-central-backend and omics-central-docker) and bookdown repos (omics-central-learn and omics-central-contribute).

The purpose of this book is not to copy, paste other works but to link works by different authors in one place and explain concepts through the lens of an omics researcher.

Chapter 2

Introduction

You can label chapter and section titles using `{#label}` after them, e.g., we can reference Chapter 2. If you do not manually label them, there will be automatic labels anyway, e.g., Chapter ??.

Figures and tables with captions will be placed in `figure` and `table` environments, respectively.

```
par(mar = c(4, 4, .1, .1))  
plot(pressure, type = 'b', pch = 19)
```

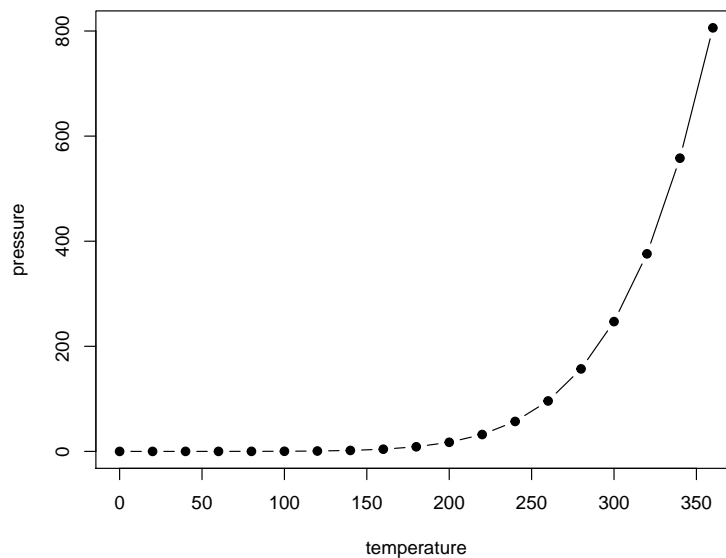


Figure 2.1: Here is a nice figure!

Table 2.1: Here is a nice table!

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa
5.1	3.5	1.4	0.3	setosa
5.7	3.8	1.7	0.3	setosa
5.1	3.8	1.5	0.3	setosa

Reference a figure by its code chunk label with the `fig:` prefix, e.g., see Figure 2.1. Similarly, you can reference tables generated from `knitr::kable()`, e.g., see Table 2.1.

```
knitr::kable(
  head(iris, 20), caption = 'Here is a nice table!',
  booktabs = TRUE
)
```

You can write citations, too. For example, we are using the **bookdown** package (Xie, 2019) in this sample book, which was built on top of R Markdown and **knitr** (Xie, 2015).

Chapter 3

Data-types

3.1 Microarrays

3.2 RNA sequencing

3.3 Nanostring

3.4 Biocrates

3.5 Multiple Reaction Monitoring

Chapter 4

Exploratory Data Analysis

4.1 Principal Component Analysis

4.1.1 Method

4.1.2 Visualizations

4.1.2.1 Scree plot

4.1.2.2 Component plot

4.1.2.3 Correlation circle

4.1.2.4 Biplot

4.2 K-Means

4.3 Hierarchical clustering

4.4 Sample plots

4.4.1 Sample correlation heatmap

4.4.2 Sample histograms

4.5 Variable plots

Chapter 5

Batch Correction

5.1 ComBat

5.2 Surrogate Variable Analysis

5.3 Model adjustment

Chapter 6

Differential Expression Analysis

6.1 Ordinary Least Squares

6.2 Linear Models for MicroArrays and RNA-Seq

6.2.1 Robust LIMMA

6.2.2 LIMMA VOOM

6.3 Significance Analysis for Microarrays (SAM)

6.4 cell-specific Analysis for Microarrays (csSAM)

Chapter 7

Network Analysis

7.1 DINGO

7.2 WGCNA

7.3 PANDA

7.4 BioNetStat

Chapter 8

Data Integration

8.1 Supervised

8.1.1 DIABLO (SGCCDA)

8.1.2 Ensemble of glmnet classifiers

8.1.3 DIABLO2 (sMB-PLSDA)

8.2 Unsupervised

8.2.1 PANDA

8.2.2 MOFA

8.2.3 JIVE

8.2.4 SNF

Chapter 9

Biological Enrichment

9.1 Enrichr

9.2 SEAR

9.3 CAMERA

9.4 Network-based Gene Set Analysis

Chapter 10

Literature Mining

Bibliography

Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.

Xie, Y. (2019). *bookdown: Authoring Books and Technical Documents with R Markdown*. R package version 0.16.