

Temas Selectos de Análisis Numérico y  
Computación Científica: Computo científico para  
el análisis de datos

Haydeé Peruyero

2024-01-24



# Contents

<b>1</b>	<b>Temas Selectos de Análisis Numérico y Computación Científica: Computo científico para el análisis de datos</b>	<b>5</b>
1.1	Temario . . . . .	5
1.2	Referencias . . . . .	6
1.3	Material interesante . . . . .	6
1.4	DataCamp . . . . .	6
<b>2</b>	<b>Shell</b>	<b>9</b>
2.1	Manipulación de archivos y directorios . . . . .	9
2.2	Manipulación de datos . . . . .	9
2.3	Tuberías y filtros . . . . .	9
2.4	Ciclos . . . . .	9
2.5	Scripts . . . . .	9
2.6	Descarga y limpieza de bases de datos . . . . .	9
<b>3</b>	<b>Git y Github</b>	<b>11</b>
3.1	Repositorios . . . . .	11
3.2	Flujo de trabajo en Git . . . . .	11
3.3	Comparando cambios . . . . .	11
3.4	Crear Ramas . . . . .	11
3.5	Actualizando ramas . . . . .	11
3.6	Revertir cambios . . . . .	11
3.7	Resolver conflictos . . . . .	11

<b>4</b>	<b>Python</b>	<b>13</b>
4.1	Tipos de datos . . . . .	13
4.2	Flujo de control . . . . .	13
4.3	Visualización de datos . . . . .	13
4.4	Manipulación de bases de datos . . . . .	13
4.5	Análisis exploratorio de bases de datos . . . . .	13
4.6	Funciones y scripts . . . . .	13
4.7	Buenas practicas . . . . .	13
4.8	Procesamiento de alto rendimiento . . . . .	13
4.9	Programación en paralelo . . . . .	13
<b>5</b>	<b>SQL</b>	<b>15</b>
5.1	Bases de datos y manipulación . . . . .	15
5.2	Explorar datos categóricos y texto no estructurado . . . . .	15
5.3	Comparación con los otros programas . . . . .	15
5.4	Valores faltantes . . . . .	15
5.5	Combinar bases de datos . . . . .	15
<b>6</b>	<b>Power BI</b>	<b>17</b>
6.1	Introducción a Power BI . . . . .	17
6.2	Transformando y visualizando datos . . . . .	17
6.3	Manipulación de bases de datos . . . . .	17
6.4	Análisis exploratorio de bases de datos . . . . .	17
6.5	Variables categóricas y continuas . . . . .	17
<b>7</b>	<b>R</b>	<b>19</b>
7.1	Tipos de datos . . . . .	19
7.2	Manipulación de bases de datos . . . . .	19
7.3	Análisis exploratorio de bases de datos . . . . .	19
7.4	Reportes con RMarkdown . . . . .	19
7.5	Páginas web . . . . .	19

## Chapter 1

# Temas Selectos de Análisis Numérico y Computación Científica: Computo científico para el análisis de datos

Curso del posgrado conjunto en Ciencias Matemáticas PCCM UNAM UMICH  
2024-2

### 1.1 Temario

1. Git y Github
2. Shell
3. Python
4. SQL
5. Power BI
6. R
7. Estadística multivariada
8. Análisis de regresión

## 1.2 Referencias

- [1] Arnold, Jeremy. Learning Microsoft Power BI, O'Reilly Media, Inc.
- [2] Beaulieu, Alan. Learning SQL, O'Reilly Media, Inc., 2020
- [3] Bruce, Peter, Bruce, Andrew and Gedeck, Peter. Practical Statistics for Data Scientists, O'Reilly Media, Inc., 2020.
- [4] Crawley, Michael J. The R book. John Wiley & Sons, 2012.
- [5] McKinney, Wes. Python for data analysis. O'Reilly Media, Inc., 2022.
- [6] Nelli, Fabio. Python Data Analytics, Apress.
- [7] Wade, Ryan. Advanced Analytics in Power BI with R and Python, Apress.
- [8] Wickham, Hadley, and Garrett Grolmund. R for data science: import, tidy, transform, visualize, and model data. O Reilly Media, Inc., 2016.
- [9] Zamora Saiz, Alfonso, et al. An Introduction to Data Analysis in R: Hands-on Coding, Data Mining, Visualization and Statistics from Scratch., Springer (2020).
- [10] Software Carpentry, The Unix Shell, <https://swcarpentry.github.io/shell-novice/>

## 1.3 Material interesante

- Bookdown.
- Software Carpentry.
- Git
- Why Git
- R Markdown Cookbook
- STHDA
- YaRrr! The Pirate's Guide to R
- Learn ggplot2 Using Shiny App
- Ggplot2: Elegant Graphics for Data Analysis
  - Versión online
- Use R! Colección Springer
- Lattice: Multivariate Data Visualization with R
- R Graphics cookbook

## 1.4 DataCamp



Figure 1.1: DataCamp





## Chapter 2

# Shell

2.1 Manipulación de archivos y directorios

2.2 Manipulación de datos

2.3 Tuberías y filtros

2.4 Ciclos

2.5 Scripts

2.6 Descarga y limpieza de bases de datos



## Chapter 3

# Git y Github

### 3.1 Repositorios

### 3.2 Flujo de trabajo en Git

### 3.3 Comparando cambios

### 3.4 Crear Ramas

### 3.5 Actualizando ramas

### 3.6 Revertir cambios

### 3.7 Resolver conflictos



## Chapter 4

# Python

- 4.1 Tipos de datos
- 4.2 Flujo de control
- 4.3 Visualización de datos
- 4.4 Manipulación de bases de datos
- 4.5 Análisis exploratorio de bases de datos
- 4.6 Funciones y scripts
- 4.7 Buenas practicas
- 4.8 Procesamiento de alto rendimiento
- 4.9 Programación en paralelo



## Chapter 5

# SQL

5.1 Bases de datos y manipulación

5.2 Explorar datos categóricos y texto no estructurado

5.3 Comparación con los otros programas

5.4 Valores faltantes

5.5 Combinar bases de datos





## Chapter 6

# Power BI

6.1 Introducción a Power BI

6.2 Transformando y visualizando datos

6.3 Manipulación de bases de datos

6.4 Análisis exploratorio de bases de datos

6.5 Variables categóricas y continuas



## Chapter 7

# R

7.1 Tipos de datos

7.2 Manipulación de bases de datos

7.3 Análisis exploratorio de bases de datos

7.4 Reportes con RMarkdown

7.5 Páginas web