1. Record Nr. UNINA9911004803403321

Autore	Campesato Oswald
Titolo	Python Tools for Data Scientists Pocket Primer
Pubbl/distr/stampa	Bloomfield : , : Mercury Learning & Information, , 2022 ©2022
ISBN	9781683928218 1683928210 9781683928225 1683928229
Edizione	[1st ed.]
Descrizione fisica	1 online resource (323 pages)
Collana	Computing
Disciplina	001.642
Soggetti	Computer programming Electronic data processing Python (Computer program language)
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Cover -- Half-Title -- Title -- Copyright -- Dedication -- Contents -- Preface -- Chapter 1: Introduction to Python -- Tools for Python -- easy_install and pip -- virtualenv -- Python Installation -- Setting the PATH Environment Variable (Windows Only) -- Launching Python on Your Machine -- The Python Interactive Interpreter -- Python Identifiers -- Lines, Indentations, and Multi-Lines -- Quotation and Comments in Python -- Saving Your Code in a Module -- Some Standard Modules in Python -- The help() and dir() Functions -- Compile Time and Runtime Code Checking -- Simple Data Types in Python -- Working with Numbers -- Working with Other Bases -- The chr() Function -- The round() Function in Python -- Formatting Numbers in Python -- Unicode and UTF-8 -- Working with Unicode -- Listing 1.1: Unicode1.py -- Working with Strings -- Comparing Strings -- Listing 1.2: Compare.py -- Formatting Strings in Python -- Uninitialized Variables and the Value None in Python -- Slicing and Splicing Strings -- Testing for Digits and Alphabetic Characters -- Listing 1.3: CharTypes.py -- Search and Replace a String in Other Strings -- Listing 1.4: FindPos1.py -- Listing 1.5: Replace1.py -- Remove Leading and Trailing Characters

pandas_quarterly_df1.py -- Data Manipulation with Pandas Data Frames (2) -- Listing 3.17: pandas_quarterly_df2.py -- Data Manipulation with Pandas Data Frames (3) -- Listing 3.18: pandas_quarterly_df3.py -- Pandas Data Frames and CSV Files -- Listing 3.19: weather_data.py -- Listing 3.20: people.csv -- Listing 3.21: people_pandas.py -- Managing Columns in Data Frames -- Switching Columns -- Appending Columns -- Deleting Columns -- Inserting Columns -- Scaling Numeric Columns -- Listing 3.22: numbers.csv.
Listing 3.23: scale_columns.py -- Managing Rows in Pandas -- Selecting a Range of Rows in Pandas -- Listing 3.24: duplicates.csv -- Listing 3.25: row_range.py -- Finding Duplicate Rows in Pandas -- Listing 3.26: duplicates.py -- Listing 3.27: drop_duplicates.py -- Inserting New Rows in Pandas -- Listing 3.28: emp_ages.csv -- Listing 3.29: insert_row.py -- Handling Missing Data in Pandas -- Listing 3.30: employees2.csv -- Listing 3.31: missing_values.py -- Multiple Types of Missing Values -- Listing 3.32: employees3.csv -- Listing 3.33: missing_multiple_types.py -- Test for Numeric Values in a Column -- Listing 3.34: test_for_numeric.py -- Replacing NaN Values in Pandas -- Listing 3.35: missing_fill_drop.py -- Sorting Data Frames in Pandas -- Listing 3.36: sort_df.py -- Working with groupby() in Pandas -- Listing 3.37: groupby1.py -- Working with apply() and mapapply() in Pandas -- Listing 3.38: apply1.py -- Listing 3.39: apply2.py -- Listing 3.40: mapapply1.py -- Listing 3.41: mapapply2.py -- Handling Outliers in Pandas -- Listing 3.42: outliers_zscores.py -- Pandas Data Frames and Scatterplots -- Listing 3.43: pandas_scatter_df.py -- Pandas Data Frames and Simple Statistics -- Listing 3.44: housing.csv -- Listing 3.45: housing_stats.py -- Aggregate Operations in Pandas Data Frames -- Listing 3.46: aggregate1.py -- Aggregate Operations with the titanic.csv Dataset -- Listing 3.47: aggregate2.py -- Save Data Frames as CSV Files and Zip Files -- Listing 3.48: save2csv.py -- Pandas Data Frames and Excel Spreadsheets -- Listing 3.49: write_people_xlsx.py -- Listing 3.50: read_people_xslx.py -- Working with JSON-based Data -- Python Dictionary and JSON -- Listing 3.51: dict2json.py -- Python, Pandas, and JSON -- Listing 3.52: pd_python_json.py -- Useful One-line Commands in Pandas -- What is Method Chaining? -- Pandas and Method Chaining.
Pandas Profiling -- Listing 3.53: titanic.csv -- Listing 3.54: profile_titanic.py -- Summary -- Chapter 4: Working with Sklearn and Scipy -- What is Sklearn? -- Sklearn Features -- The Digits Dataset in Sklearn -- Listing 4.1: load_digits1.py -- Listing 4.2: load_digits2.py -- Listing 4.3: sklearn_digits.py -- The train_test_split() Class in Sklearn -- Selecting Columns for X and y -- What is Feature Engineering? -- The Iris Dataset in Sklearn (1) -- Listing 4.4: sklearn_iris1.py -- Sklearn, Pandas, and the Iris Dataset -- Listing 4.5: pandas_iris.py -- The Iris Dataset in Sklearn (2) -- Listing 4.6: sklearn_iris2.py -- The Faces Dataset in Sklearn (Optional) -- Listing 4.7: sklearn_faces.py -- What is SciPy? -- Installing SciPy -- Permutations and Combinations in SciPy -- Listing 4.8: scipy_perms.py -- Listing 4.9: scipy_combinatorics.py -- Calculating Log Sums -- Listing 4.10: scipy_matrix_inv.py -- Calculating Polynomial Values -- Listing 4.11: scipy_poly.py -- Calculating the Determinant of a Square Matrix -- Listing 4.12: scipy_determinant.py -- Calculating the Inverse of a Matrix -- Listing 4.13: scipy_matrix_inv.py -- Calculating Eigenvalues and Eigenvectors -- Listing 4.14: scipy_eigen.py -- Calculating Integrals (Calculus) -- Listing 4.15: scipy_integrate.py -- Calculating Fourier Transforms -- Listing 4.16: scipy_fourier.py -- Flipping Images in SciPy -- Listing 4.17: scipy_flip_image.py -- Rotating Images in SciPy

	-- Listing 4.18: scipy_rotate_image.py -- Google Colaboratory -- Uploading CSV Files in Google Colaboratory -- Listing 4.19: upload_csv_file.ipynb -- Summary -- Chapter 5: Data Cleaning Tasks -- What is Data Cleaning? -- Data Cleaning for Personal Titles -- Data Cleaning in SQL -- Replace NULL with 0 -- Replace NULL Values with the Average Value -- Listing 5.1: replace_null_values.sql. Replace Multiple Values with a Single Value.
Sommario/riassunto	"This book contains a fast-paced introduction to as much relevant information about Python tools for data scientists as possible that can be reasonably included in a book of this size. If you are a novice, this book will give you a starting point from which you can decide which Python technologies that you want to explore in greater detail. You will be exposed to features of Numpy and Pandas, how to write regular expressions, and how to perform data cleaning tasks. Some topics are presented in a cursory manner, which is for two main reasons. First, it's important that you be exposed to these concepts. In some cases you will find topics that might pique your interest, and hence motivate you to learn more about them through self-study; in other cases you will probably be satisfied with a brief introduction. In other words, you will decide whether or not to delve into more detail regarding each of the topics in this book"--