

Module-03, Python for Data Analysis

Data Exploration (Pandas-Exercise)

Dostdar Ali
Instructor

Data science and Artificial Intelligence
3-Months Course
at
Karakaroum international Univrsity

January 4, 2024



Table of Contents

1 SF Salaries Dataset from Kaggle

2 Ecommerce Purchases Dataset



SF Salaries exercise

The data have 148654 entries and 13 columns, like `Id`, `EmployeeName`, `JobTitle`, `BasePay`, `OvertimePay`, `OtherPay`, `Benefits`, `TotalPay`, `TotalPayBenefits`, `Year`, `Notes`, `Agency`, `Status`. Now exercise for you to practice your pandas skills! We will be using the SF Salaries Dataset from Kaggle! Just follow along and complete the tasks. The tasks will get harder and harder as you go along.

	Id	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefits	Year	Notes	Agency	Status
0	1	NATHANIEL FORD	GENERAL MANAGER-METROPOLITAN TRANSIT AUTHORITY	167411.18	0.0	400184.25	NaN	567595.43	567595.43	2011	NaN	San Francisco	NaN

- Import pandas as `pd`.
- Read `Salaries.csv` as a dataframe called `sal`
- Check the head of the DataFrame.
- Use the `.info()` method to find out how many entries there are.



SF Salaries exercise

The data have 148654 entries and 13 columns, like `Id`, `EmployeeName`, `JobTitle`, `BasePay`, `OvertimePay`, `OtherPay`, `Benefits`, `TotalPay`, `TotalPayBenefits`, `Year`, `Notes`, `Agency`, `Status`. Now exercise for you to practice your pandas skills! We will be using the SF Salaries Dataset from Kaggle! Just follow along and complete the tasks. The tasks will get harder and harder as you go along.

	Id	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefits	Year	Notes	Agency	Status
0	1	NATHANIEL FORD	GENERAL MANAGER-METROPOLITAN TRANSIT AUTHORITY	167411.18	0.0	400184.25	NaN	567595.43	567595.43	2011	NaN	San Francisco	NaN

- Import pandas as pd.
- Read Salaries.csv as a dataframe called sal
- Check the head of the DataFrame.
- Use the `.info()` method to find out how many entries there are.



SF Salaries exercise

The data have 148654 entries and 13 columns, like `Id`, `EmployeeName`, `JobTitle`, `BasePay`, `OvertimePay`, `OtherPay`, `Benefits`, `TotalPay`, `TotalPayBenefits`, `Year`, `Notes`, `Agency`, `Status`. Now exercise for you to practice your pandas skills! We will be using the SF Salaries Dataset from Kaggle! Just follow along and complete the tasks. The tasks will get harder and harder as you go along.

	Id	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefits	Year	Notes	Agency	Status
0	1	NATHANIEL FORD	GENERAL MANAGER-METROPOLITAN TRANSIT AUTHORITY	167411.18	0.0	400184.25	NaN	567595.43	567595.43	2011	NaN	San Francisco	NaN

- Import pandas as `pd`.
- Read `Salaries.csv` as a dataframe called `sal`
- Check the head of the DataFrame.
- Use the `.info()` method to find out how many entries there are.



SF Salaries exercise

The data have 148654 entries and 13 columns, like `Id`, `EmployeeName`, `JobTitle`, `BasePay`, `OvertimePay`, `OtherPay`, `Benefits`, `TotalPay`, `TotalPayBenefits`, `Year`, `Notes`, `Agency`, `Status`. Now exercise for you to practice your pandas skills! We will be using the SF Salaries Dataset from Kaggle! Just follow along and complete the tasks. The tasks will get harder and harder as you go along.

	Id	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefits	Year	Notes	Agency	Status
0	1	NATHANIEL FORD	GENERAL MANAGER- METROPOLITAN TRANSIT AUTHORITY	167411.18	0.0	400184.25	NaN	567595.43	567595.43	2011	NaN	San Francisco	NaN

- Import pandas as `pd`.
- Read `Salaries.csv` as a dataframe called `sal`
- Check the head of the DataFrame.
- Use the `.info()` method to find out how many entries there are.



- What is the average BasePay?
- What is the highest amount of OvertimePay in the dataset?
- What is the job title of JOSEPH DRISCOLL ? Note: Use all caps, otherwise you may get an answer that doesn't match up (there is also a lowercase Joseph Driscoll).
- How much does JOSEPH DRISCOLL make (including benefits)?
- What is the name of highest paid person (including benefits)?
- What is the name of lowest paid person (including benefits)? Do you notice something strange about how much he or she is paid?
- What was the average (mean) BasePay of all employees per year? (2011-2014)?
- How many unique job titles are there?
- What are the top 5 most common jobs?
- How many Job Titles were represented by only one person in 2013? (e.g. Job Titles with only one occurrence in 2013?)



- What is the average BasePay?
- What is the highest amount of OvertimePay in the dataset?
- What is the job title of JOSEPH DRISCOLL ? Note: Use all caps, otherwise you may get an answer that doesn't match up (there is also a lowercase Joseph Driscoll).
- How much does JOSEPH DRISCOLL make (including benefits)?
- What is the name of highest paid person (including benefits)?
- What is the name of lowest paid person (including benefits)? Do you notice something strange about how much he or she is paid?
- What was the average (mean) BasePay of all employees per year? (2011-2014)?
- How many unique job titles are there?
- What are the top 5 most common jobs?
- How many Job Titles were represented by only one person in 2013? (e.g. Job Titles with only one occurrence in 2013?)



- What is the average BasePay?
- What is the highest amount of OvertimePay in the dataset?
- What is the job title of JOSEPH DRISCOLL ? Note: Use all caps, otherwise you may get an answer that doesn't match up (there is also a lowercase Joseph Driscoll).
- How much does JOSEPH DRISCOLL make (including benefits)?
- What is the name of highest paid person (including benefits)?
- What is the name of lowest paid person (including benefits)? Do you notice something strange about how much he or she is paid?
- What was the average (mean) BasePay of all employees per year? (2011-2014)?
- How many unique job titles are there?
- What are the top 5 most common jobs?
- How many Job Titles were represented by only one person in 2013? (e.g. Job Titles with only one occurrence in 2013?)



- What is the average BasePay?
- What is the highest amount of OvertimePay in the dataset?
- What is the job title of JOSEPH DRISCOLL ? Note: Use all caps, otherwise you may get an answer that doesn't match up (there is also a lowercase Joseph Driscoll).
- How much does JOSEPH DRISCOLL make (including benefits)?
- What is the name of highest paid person (including benefits)?
- What is the name of lowest paid person (including benefits)? Do you notice something strange about how much he or she is paid?
- What was the average (mean) BasePay of all employees per year? (2011-2014)?
- How many unique job titles are there?
- What are the top 5 most common jobs?
- How many Job Titles were represented by only one person in 2013? (e.g. Job Titles with only one occurrence in 2013?)



- What is the average BasePay?
- What is the highest amount of OvertimePay in the dataset?
- What is the job title of JOSEPH DRISCOLL ? Note: Use all caps, otherwise you may get an answer that doesn't match up (there is also a lowercase Joseph Driscoll).
- How much does JOSEPH DRISCOLL make (including benefits)?
- What is the name of highest paid person (including benefits)?
- What is the name of lowest paid person (including benefits)? Do you notice something strange about how much he or she is paid?
- What was the average (mean) BasePay of all employees per year? (2011-2014)?
- How many unique job titles are there?
- What are the top 5 most common jobs?
- How many Job Titles were represented by only one person in 2013? (e.g. Job Titles with only one occurrence in 2013?)



- What is the average BasePay?
- What is the highest amount of OvertimePay in the dataset?
- What is the job title of JOSEPH DRISCOLL ? Note: Use all caps, otherwise you may get an answer that doesn't match up (there is also a lowercase Joseph Driscoll).
- How much does JOSEPH DRISCOLL make (including benefits)?
- What is the name of highest paid person (including benefits)?
- What is the name of lowest paid person (including benefits)? Do you notice something strange about how much he or she is paid?
- What was the average (mean) BasePay of all employees per year? (2011-2014)?
- How many unique job titles are there?
- What are the top 5 most common jobs?
- How many Job Titles were represented by only one person in 2013? (e.g. Job Titles with only one occurrence in 2013?)



- What is the average BasePay?
- What is the highest amount of OvertimePay in the dataset?
- What is the job title of JOSEPH DRISCOLL ? Note: Use all caps, otherwise you may get an answer that doesn't match up (there is also a lowercase Joseph Driscoll).
- How much does JOSEPH DRISCOLL make (including benefits)?
- What is the name of highest paid person (including benefits)?
- What is the name of lowest paid person (including benefits)? Do you notice something strange about how much he or she is paid?
- What was the average (mean) BasePay of all employees per year? (2011-2014)?
- How many unique job titles are there?
- What are the top 5 most common jobs?
- How many Job Titles were represented by only one person in 2013? (e.g. Job Titles with only one occurrence in 2013?)



- What is the average BasePay?
- What is the highest amount of OvertimePay in the dataset?
- What is the job title of JOSEPH DRISCOLL ? Note: Use all caps, otherwise you may get an answer that doesn't match up (there is also a lowercase Joseph Driscoll).
- How much does JOSEPH DRISCOLL make (including benefits)?
- What is the name of highest paid person (including benefits)?
- What is the name of lowest paid person (including benefits)? Do you notice something strange about how much he or she is paid?
- What was the average (mean) BasePay of all employees per year? (2011-2014)?
- How many unique job titles are there?
- What are the top 5 most common jobs?
- How many Job Titles were represented by only one person in 2013? (e.g. Job Titles with only one occurrence in 2013?)



- What is the average BasePay?
- What is the highest amount of OvertimePay in the dataset?
- What is the job title of JOSEPH DRISCOLL ? Note: Use all caps, otherwise you may get an answer that doesn't match up (there is also a lowercase Joseph Driscoll).
- How much does JOSEPH DRISCOLL make (including benefits)?
- What is the name of highest paid person (including benefits)?
- What is the name of lowest paid person (including benefits)? Do you notice something strange about how much he or she is paid?
- What was the average (mean) BasePay of all employees per year? (2011-2014)?
- How many unique job titles are there?
- What are the top 5 most common jobs?
- How many Job Titles were represented by only one person in 2013? (e.g. Job Titles with only one occurrence in 2013?)



- What is the average BasePay?
- What is the highest amount of OvertimePay in the dataset?
- What is the job title of JOSEPH DRISCOLL ? Note: Use all caps, otherwise you may get an answer that doesn't match up (there is also a lowercase Joseph Driscoll).
- How much does JOSEPH DRISCOLL make (including benefits)?
- What is the name of highest paid person (including benefits)?
- What is the name of lowest paid person (including benefits)? Do you notice something strange about how much he or she is paid?
- What was the average (mean) BasePay of all employees per year? (2011-2014)?
- How many unique job titles are there?
- What are the top 5 most common jobs?
- How many Job Titles were represented by only one person in 2013? (e.g. Job Titles with only one occurrence in 2013?)



Ecommerce Purchases

In this Exercise you will be given some Fake Data about some purchases done through Amazon! which have 10000 entries and 14 columns **Address**, **Lot**, **AM to PM**, **Browser Info**, **Company**, **Credit Card**, **CC Exp Date**, **CC Security Code**, **CC Provider**, **Email**, **Job**, **IP Address**, **Language**, **Purchase Price**. Just go ahead and follow the directions and try your best to answer the questions and complete the tasks. Feel free to reference the solutions. Most of the tasks can be solved in different ways. For the most part, the questions get progressively harder. Please excuse anything that doesn't make "Real-World" sense in the dataframe, all the data is fake and made-up. Also note that all of these questions can be answered with one line of code.

Address	Lot	AM to PM	Browser Info	Company	Credit Card	CC Exp Date	CC Security Code	CC Provider	Email	Job	IP Address	Language	Purchase Price
e Camp Apt.orough, NE 77...	46	PM	Opera/9.56.(X11; Linux x86_64; sl-SI) Presto/2...	Martinez-Herman	6011929061123406	02/20	900	JCB 16 digit	pduinlap@yahoo.com	Scientist, product/process development	149.146.147.205	el	98



- Import pandas and read in the Ecommerce Purchases csv file and set it to a DataFrame called ecom.
- Check the head of the DataFrame. How many rows and columns are there?
- What is the average Purchase Price? What were the highest and lowest purchase prices?
- How many people have English 'en' as their Language of choice on the website?
- How many people have the job title of "Lawyer" ?
- How many people made the purchase during the AM and how many people made the purchase during PM ?
- What are the 5 most common Job Titles?
- Someone made a purchase that came from Lot: "90 WT" , what was the Purchase Price for this transaction?



- Import pandas and read in the Ecommerce Purchases csv file and set it to a DataFrame called ecom.
- Check the head of the DataFrame. How many rows and columns are there?
- What is the average Purchase Price? What were the highest and lowest purchase prices?
- How many people have English 'en' as their Language of choice on the website?
- How many people have the job title of "Lawyer" ?
- How many people made the purchase during the AM and how many people made the purchase during PM ?
- What are the 5 most common Job Titles?
- Someone made a purchase that came from Lot: "90 WT" , what was the Purchase Price for this transaction?



- Import pandas and read in the Ecommerce Purchases csv file and set it to a DataFrame called ecom.
- Check the head of the DataFrame. How many rows and columns are there?
- What is the average Purchase Price? What were the highest and lowest purchase prices?
- How many people have English 'en' as their Language of choice on the website?
- How many people have the job title of "Lawyer" ?
- How many people made the purchase during the AM and how many people made the purchase during PM ?
- What are the 5 most common Job Titles?
- Someone made a purchase that came from Lot: "90 WT" , what was the Purchase Price for this transaction?



- Import pandas and read in the Ecommerce Purchases csv file and set it to a DataFrame called ecom.
- Check the head of the DataFrame. How many rows and columns are there?
- What is the average Purchase Price? What were the highest and lowest purchase prices?
- How many people have English 'en' as their Language of choice on the website?
- How many people have the job title of "Lawyer" ?
- How many people made the purchase during the AM and how many people made the purchase during PM ?
- What are the 5 most common Job Titles?
- Someone made a purchase that came from Lot: "90 WT" , what was the Purchase Price for this transaction?



- Import pandas and read in the Ecommerce Purchases csv file and set it to a DataFrame called ecom.
- Check the head of the DataFrame. How many rows and columns are there?
- What is the average Purchase Price? What were the highest and lowest purchase prices?
- How many people have English 'en' as their Language of choice on the website?
- How many people have the job title of "Lawyer" ?
- How many people made the purchase during the AM and how many people made the purchase during PM ?
- What are the 5 most common Job Titles?
- Someone made a purchase that came from Lot: "90 WT" , what was the Purchase Price for this transaction?



- Import pandas and read in the Ecommerce Purchases csv file and set it to a DataFrame called ecom.
- Check the head of the DataFrame. How many rows and columns are there?
- What is the average Purchase Price? What were the highest and lowest purchase prices?
- How many people have English 'en' as their Language of choice on the website?
- How many people have the job title of "Lawyer" ?
- How many people made the purchase during the AM and how many people made the purchase during PM ?
- What are the 5 most common Job Titles?
- Someone made a purchase that came from Lot: "90 WT" , what was the Purchase Price for this transaction?



- Import pandas and read in the Ecommerce Purchases csv file and set it to a DataFrame called ecom.
- Check the head of the DataFrame. How many rows and columns are there?
- What is the average Purchase Price? What were the highest and lowest purchase prices?
- How many people have English 'en' as their Language of choice on the website?
- How many people have the job title of "Lawyer" ?
- How many people made the purchase during the AM and how many people made the purchase during PM ?
- What are the 5 most common Job Titles?
- Someone made a purchase that came from Lot: "90 WT" , what was the Purchase Price for this transaction?



- Import pandas and read in the Ecommerce Purchases csv file and set it to a DataFrame called ecom.
- Check the head of the DataFrame. How many rows and columns are there?
- What is the average Purchase Price? What were the highest and lowest purchase prices?
- How many people have English 'en' as their Language of choice on the website?
- How many people have the job title of "Lawyer" ?
- How many people made the purchase during the AM and how many people made the purchase during PM ?
- What are the 5 most common Job Titles?
- Someone made a purchase that came from Lot: "90 WT" , what was the Purchase Price for this transaction?



- What is the email of the person with the following Credit Card Number: 4926535242672853
- How many people have American Express as their Credit Card Provider and made a purchase above 95 ?
- Hard: How many people have a credit card that expires in 2025?
- Hard: What are the top 5 most popular email providers/hosts (e.g. gmail.com, yahoo.com, etc...)



- What is the email of the person with the following Credit Card Number: 4926535242672853
- How many people have American Express as their Credit Card Provider and made a purchase above 95 ?
- Hard: How many people have a credit card that expires in 2025?
- Hard: What are the top 5 most popular email providers/hosts (e.g. gmail.com, yahoo.com, etc...)



- What is the email of the person with the following Credit Card Number: 4926535242672853
- How many people have American Express as their Credit Card Provider and made a purchase above 95 ?
- Hard: How many people have a credit card that expires in 2025?
- Hard: What are the top 5 most popular email providers/hosts (e.g. gmail.com, yahoo.com, etc...)



- What is the email of the person with the following Credit Card Number: 4926535242672853
- How many people have American Express as their Credit Card Provider and made a purchase above 95 ?
- Hard: How many people have a credit card that expires in 2025?
- Hard: What are the top 5 most popular email providers/hosts (e.g. gmail.com, yahoo.com, etc...)



Great Job
Thank yo

