

Workshop 3 | Pandas review

Master's Degree in Data Sciences | EAFIT University

Course: Fundamentals of Data Science

Group: Sara Martinez Rendon (smartiner4@eafit.edu.co) | Heider Zapata (hzapata1@eafit.edu.co) | Yeison Londoño (ylondon4@eafit.edu.co)

Date: 31/01/2026

```
In [ ]: # Library
import pandas as pd
import random as rand
```

```
In [ ]: # 1. select path
from google.colab import drive
drive.mount('/content/drive')
ruta_data = '/content/drive/MyDrive/Colab Notebooks/Data/grupo1626.xlsx'
ruta_data_2 = '/content/drive/MyDrive/Colab Notebooks/Data/PC - 1626.xlsx'
```

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).

```
In [ ]: # 2. Import data
grupo = pd.read_excel(ruta_data)
carreras = pd.read_excel(ruta_data_2, sheet_name="Hoja2")
grupo.head()
```

	Username	Last Name	First Name	Email
0	amcastroy	CASTRO YEPES	ANA MARIA	amcastroy@eafit.brightspace.com
1	ausmec	USME CARDONA	AINARA	ausmec@eafit.brightspace.com
2	clbruckner	BRUCKNER RIBEIRO	CHRISTIAN	clbruckner@eafit.brightspace.com
3	dacarmonm1	CARMONA MONTES	DAVID ALEJANDRO	dacarmonm1@eafit.brightspace.com
4	dsanchezu	SANCHEZ URREA	DANIEL	dsanchezu@eafit.brightspace.com

```
In [ ]: carreras.head()
```

Out[]:

	Username	carrera
0	amcastroy	Ingeniería de Producción - Ingeniería de Produ...
1	ausmec	Ingeniería Mecánica - Ingeniería Mecánica-20241
2	clbruckner	Ingeniería Civil - Ingeniería Civil-20242
3	dacarmonm1	Administración de Negocios - Administración Ne...
4	dsanchezu	Biología - Biología-20261

In []:

```
# 3. Explore student's names with group
```

```
nombres = grupo["First Name"]  
nombres
```

Out[]:

	First Name
0	ANA MARIA
1	AINARA
2	CHRISTIAN
3	DAVID ALEJANDRO
4	DANIEL
5	EMILIANA
6	EVA MARIA
7	EMMANUEL
8	ELIZABETH SAYURI
9	JUAN
10	JUAN LUCAS
11	JUAN SEBASTIAN
12	LAURA
13	MANUELA
14	MIGUEL ANGEL
15	MARIANA
16	NICOLLE
17	Oscar Alonso
18	PAULA
19	SOFIA
20	SOFIA
21	SUSANA
22	VALERIA
23	VICTOR MANUEL
24	VIOLETA

dtype: object

In []: # 4. Select indexes

```
grupo.set_index('Username', inplace=True)
grupo.head()
```

Out[]:

	Last Name	First Name	Email
Username			
amcastroy	CASTRO YEPES	ANA MARIA	amcastroy@eafit.brightspace.com
ausmec	USME CARDONA	AINARA	ausmec@eafit.brightspace.com
clbruckner	BRUCKNER RIBEIRO	CHRISTIAN	clbruckner@eafit.brightspace.com
dacarmonm1	CARMONA MONTES	DAVID ALEJANDRO	dacarmonm1@eafit.brightspace.com
dsanchezu	SANCHEZ URREA	DANIEL	dsanchezu@eafit.brightspace.com

In []: `carreras.set_index('Username', inplace=True)
carreras.head()`

Out[]:

carrera

Username	
amcastroy	Ingeniería de Producción - Ingeniería de Produ...
ausmec	Ingeniería Mecánica - Ingeniería Mecánica-20241
clbruckner	Ingeniería Civil - Ingeniería Civil-20242
dacarmonm1	Administración de Negocios - Administración Ne...
dsanchezu	Biología - Biología-20261

In []: `# 4. concatenate two dataframes
final_data = pd.concat([grupo, carreras], axis=1)
final_data`

Out[]:

	Last Name	First Name	Email	carrera
Username				
amcastroy	CASTRO YEPES	ANA MARIA	amcastroy@eafit.brightspace.com	Ingeniería de Producción - Ingeniería de Produ...
ausmec	USME CARDONA	AINARA	ausmec@eafit.brightspace.com	Ingeniería Mecánica - Ingeniería Mecánica-20241
clbruckner	BRUCKNER RIBEIRO	CHRISTIAN	clbruckner@eafit.brightspace.com	Ingeniería Civil - Ingeniería Civil-20242
dacarmonm1	CARMONA MONTES	DAVID ALEJANDRO	dacarmonm1@eafit.brightspace.com	Administración de Negocios - Administración Ne...
dsanchezu	SANCHEZ URREA	DANIEL	dsanchezu@eafit.brightspace.com	Biología - Biología-20261
egiraldov2	GIRALDO VELEZ	EMILIANA	egiraldov2@eafit.brightspace.com	Negocios Internacionales - Negocios Internacio...
emmedranog	MEDRANO GARCES	EVA MARIA	emmedranog@eafit.brightspace.com	Ingeniería Civil - Ingeniería Civil-20242
esalazare1	SALAZAR ESTRADA	EMMANUEL	esalazare1@eafit.brightspace.com	Mercadeo - Mercadeo-20222
esrodrigue	RODRÍGUEZ ESPITIA	ELIZABETH SAYURI	esrodrigue@eafit.brightspace.com	Mercadeo - Mercadeo-20222
jcgarczonc	GARZON CANO	JUAN	jcgarczonc@eafit.brightspace.com	Ingeniería Industrial - Ingeniería Industrial-...
jlgutiers1	GUTIERREZ SALAZAR	JUAN LUCAS	jlgutiers1@eafit.brightspace.com	Negocios Internacionales - Negocios Internacio...
jsjaramilp	JARAMILLO PEREZ	JUAN SEBASTIAN	jsjaramilp@eafit.brightspace.com	Economía - Economía-20242

	Last Name	First Name	Email	carrera
Username				
lgilm	GIL MONTOYA	LAURA	lgilm@eafit.brightspace.com	Mercadeo - Mercadeo-2022
mariasc4	ARIAS CASTRILLON	MANUELA	mariasc4@eafit.brightspace.com	Administración de Negocios - Administración Ne...
maviolag	VIOLA GONZALEZ	MIGUEL ANGEL	maviolag@eafit.brightspace.com	Negocios Internacionales - Negocios Internacio...
msandovala	SANDOVAL ABDALLA	MARIANA	msandovala@eafit.brightspace.com	Derecho - Derecho-20242
ndosornoq	OSORNO QUINTERO	NICOLLE	ndosornoq@eafit.brightspace.com	Ingeniería Mecánica - Ingeniería Mecánica-20241
oarestrep	Restrepo Velez	Oscar Alonso	oarestrep@eafit.brightspace.com	Administración de Negocios - Administración Ne...
palunap	LUNA PEREZ	PAULA	palunap@eafit.brightspace.com	Biología - Biología-20261
sdelgadoq	DELGADO QUINTERO	SOFIA	sdelgadoq@eafit.brightspace.com	Biología - Biología-20261
sdurangoa	DURANGO ARANGO	SOFIA	sdurangoa@eafit.brightspace.com	Ingeniería Diseño de Producto - Ingen Diseño d...
smazob	MAZO BOTERO	SUSANA	smazob@eafit.brightspace.com	Comunicación Social - Comunicación Social-20261
vmarquezr	MÁRQUEZ ROJAS	VALERIA	vmarquezr@eafit.brightspace.com	Administración de Negocios - Administración Ne...
vmriascosr	RIASCOS RUIZ	VICTOR MANUEL	vmriascosr@eafit.brightspace.com	Mercadeo - Mercadeo-20222

	Last Name	First Name	Email	carrera
Username				
vvallejoh1	VALLEJO HURTADO	VIOLETA	vvallejoh1@eafit.brightspace.com	Administración de Negocios - Administración Ne...
In []:	# 5. verify shape final_data.shape			

Out[]: (25, 4)

In []: #6. reset index
final_data = final_data.reset_index()
final_data

Out[]:	Username	Last Name	First Name	Email	carrera
0	amcastroy	CASTRO YEPES	ANA MARIA	amcastroy@eafit.brightspace.com	Ingeniería d Producción Ingeniería d Produ
1	ausmec	USME CARDONA	AINARA	ausmec@eafit.brightspace.com	Ingenierí Mecánica Ingenierí Mecánica 2024
2	clbruckner	BRUCKNER RIBEIRO	CHRISTIAN	clbruckner@eafit.brightspace.com	Ingeniería Civ - Ingenierí Civil-2024
3	dacarmonm1	CARMONA MONTES	DAVID ALEJANDRO	dacarmonm1@eafit.brightspace.com	Administració de Negocios Administració Ne
4	dsanchezu	SANCHEZ URREA	DANIEL	dsanchezu@eafit.brightspace.com	Biología Biología-2026
5	egiraldov2	GIRALDO VELEZ	EMILIANA	egiraldov2@eafit.brightspace.com	Negocic Internacionale - Negocic Internacio
6	emmedranog	MEDRANO GARCES	EVA MARIA	emmedranog@eafit.brightspace.com	Ingeniería Civ - Ingenierí Civil-2024
7	esalazare1	SALAZAR ESTRADA	EMMANUEL	esalazare1@eafit.brightspace.com	Mercadeo Mercadec 2022
8	esrodrigue	RODRÍGUEZ ESPITIA	ELIZABETH SAYURI	esrodrigue@eafit.brightspace.com	Mercadeo Mercadec 2022
9	jcgarczonc	GARZON CANO	JUAN	jcgarczonc@eafit.brightspace.com	Ingenierí Industrial Ingenierí Industrial-
10	jl gutiers1	GUTIERREZ SALAZAR	JUAN LUCAS	jl gutiers1@eafit.brightspace.com	Negocic Internacionale - Negocic Internacio
11	jsjaramilp	JARAMILLO PEREZ	JUAN SEBASTIAN	jsjaramilp@eafit.brightspace.com	Economía Economía 2024

	Username	Last Name	First Name	Email	carrera
12	lgilm	GIL MONTOYA	LAURA	lgilm@eafit.brightspace.com	Mercadeo Mercadec 2022
13	mariasc4	ARIAS CASTRILLON	MANUELA	mariasc4@eafit.brightspace.com	Administració de Negocios Administració Ne
14	maviolag	VIOLA GONZALEZ	MIGUEL ANGEL	maviolag@eafit.brightspace.com	Negocios Internacionales - Negocios Internacion
15	msandovala	SANDOVAL ABDALLA	MARIANA	msandovala@eafit.brightspace.com	Derecho Derechc 2024
16	ndosornoq	OSORNO QUINTERO	NICOLLE	ndosornoq@eafit.brightspace.com	Ingeniería Mecánica Ingeniería Mecánica 2024
17	oarestrep	Restrepo Velez	Oscar Alonso	oarestrep@eafit.brightspace.com	Administración de Negocios Administració Ne
18	palunap	LUNA PEREZ	PAULA	palunap@eafit.brightspace.com	Biología Biología-2026
19	sdelgadoq	DELGADO QUINTERO	SOFIA	sdelgadoq@eafit.brightspace.com	Biología Biología-2026
20	sdurangoa	DURANGO ARANGO	SOFIA	sdurangoa@eafit.brightspace.com	Ingeniería Diseño de Producto Ingen Diseñ d
21	smazob	MAZO BOTERO	SUSANA	smazob@eafit.brightspace.com	Comunicación Social Comunicació Social-2026
22	vmarquezr	MÁRQUEZ ROJAS	VALERIA	vmarquezr@eafit.brightspace.com	Administración de Negocios Administració Ne
23	vmriascosr	RIASCOS RUIZ	VICTOR MANUEL	vmriascosr@eafit.brightspace.com	Mercadeo Mercadec 2022

Username	Last Name	First Name	Email	carrera
24 vvallejoh1	VALLEJO HURTADO	VIOLETA	vvallejoh1@eafit.brightspace.com	Administració de Negocios Administració Ne

```
In [ ]: # 7 add final grade

grade_rand = {'final_grade':[rand.randint(0,5) for i in range(len(grupo))]} 

#convertir el diccionario a df
df_grade = pd.DataFrame(grade_rand)
df_grade
```

Out[]: final_grade

	final_grade
0	5
1	0
2	0
3	3
4	3
5	1
6	4
7	1
8	2
9	5
10	4
11	0
12	4
13	3
14	2
15	5
16	1
17	4
18	0
19	0
20	3
21	4
22	2
23	3
24	4

In []: # 8. concatenar

```
final_data = pd.concat([final_data,df_grade], axis=1)  
final_data
```

Out[]:	Username	Last Name	First Name	Email	carrera
0	amcastroy	CASTRO YEPES	ANA MARIA	amcastroy@eafit.brightspace.com	Ingeniería d Producción Ingeniería d Produ
1	ausmec	USME CARDONA	AINARA	ausmec@eafit.brightspace.com	Ingenierí Mecánica Ingenierí Mecánica 2024
2	clbruckner	BRUCKNER RIBEIRO	CHRISTIAN	clbruckner@eafit.brightspace.com	Ingeniería Civ - Ingenierí Civil-2024
3	dacarmonm1	CARMONA MONTES	DAVID ALEJANDRO	dacarmonm1@eafit.brightspace.com	Administració de Negocios Administració Ne
4	dsanchezu	SANCHEZ URREA	DANIEL	dsanchezu@eafit.brightspace.com	Biología Biología-2026
5	egiraldov2	GIRALDO VELEZ	EMILIANA	egiraldov2@eafit.brightspace.com	Negocic Internacionale - Negocic Internacio
6	emmedranog	MEDRANO GARCES	EVA MARIA	emmedranog@eafit.brightspace.com	Ingeniería Civ - Ingenierí Civil-2024
7	esalazare1	SALAZAR ESTRADA	EMMANUEL	esalazare1@eafit.brightspace.com	Mercadeo Mercadec 2022
8	esrodrigue	RODRÍGUEZ ESPITIA	ELIZABETH SAYURI	esrodrigue@eafit.brightspace.com	Mercadeo Mercadec 2022
9	jcgarczonc	GARZON CANO	JUAN	jcgarczonc@eafit.brightspace.com	Ingenierí Industrial Ingenierí Industrial-
10	jl gutiers1	GUTIERREZ SALAZAR	JUAN LUCAS	jl gutiers1@eafit.brightspace.com	Negocic Internacionale - Negocic Internacio
11	jsjaramilp	JARAMILLO PEREZ	JUAN SEBASTIAN	jsjaramilp@eafit.brightspace.com	Economía Economía 2024

	Username	Last Name	First Name	Email	carrera
12	lgilm	GIL MONTOYA	LAURA	lgilm@eafit.brightspace.com	Mercadeo Mercadec 2022
13	mariasc4	ARIAS CASTRILLON	MANUELA	mariasc4@eafit.brightspace.com	Administració de Negocios Administració Ne
14	maviolag	VIOLA GONZALEZ	MIGUEL ANGEL	maviolag@eafit.brightspace.com	Negocios Internacionales - Negocios Internacion
15	msandovala	SANDOVAL ABDALLA	MARIANA	msandovala@eafit.brightspace.com	Derecho Derechc 2024
16	ndosornoq	OSORNO QUINTERO	NICOLLE	ndosornoq@eafit.brightspace.com	Ingeniería Mecánica Ingeniería Mecánica 2024
17	oarestrep	Restrepo Velez	Oscar Alonso	oarestrep@eafit.brightspace.com	Administración de Negocios Administració Ne
18	palunap	LUNA PEREZ	PAULA	palunap@eafit.brightspace.com	Biología Biología-2026
19	sdelgadoq	DELGADO QUINTERO	SOFIA	sdelgadoq@eafit.brightspace.com	Biología Biología-2026
20	sdurangoa	DURANGO ARANGO	SOFIA	sdurangoa@eafit.brightspace.com	Ingeniería Diseño de Producto Ingen Diseñ d
21	smazob	MAZO BOTERO	SUSANA	smazob@eafit.brightspace.com	Comunicación Social Comunicació Social-2026
22	vmarquezr	MÁRQUEZ ROJAS	VALERIA	vmarquezr@eafit.brightspace.com	Administración de Negocios Administració Ne
23	vmriascosr	RIASCOS RUIZ	VICTOR MANUEL	vmriascosr@eafit.brightspace.com	Mercadeo Mercadec 2022

Username	Last Name	First Name	Email	carrera
24 vvallejoh1	VALLEJO HURTADO	VIOLETA	vvallejoh1@eafit.brightspace.com	Administració de Negocios Administració Ne

In []: # 9. Nota promedio por curso

```
promedio_curso = final_data.groupby('carrera')['final_grade'].mean()
promedio_curso_sorted = promedio_curso.sort_values(ascending=False)
print(promedio_curso_sorted)
```

carrera	
Derecho - Derecho-20242	5.000000
Ingeniería Industrial - Ingeniería Industrial-20252	5.000000
Ingeniería de Producción - Ingeniería de Producción-20251	5.000000
Comunicación Social - Comunicación Social-20261	4.000000
Administración de Negocios - Administración Negocios-20222	3.200000
Ingeniería Diseño de Producto - Ingen Diseño de Producto-20252	3.000000
Mercadeo - Mercadeo-20222	2.500000
Negocios Internacionales - Negocios Internacionales-20223	2.333333
Ingeniería Civil - Ingeniería Civil-20242	2.000000
Biología - Biología-20261	1.000000
Ingeniería Mecánica - Ingeniería Mecánica-20241	0.500000
Economía - Economía-20242	0.000000
Name: final_grade, dtype: float64	