

The background is a deep blue gradient. On the left side, there are several interlocking gears of different sizes, some with a glowing effect. To the right of the gears, a complex network graph is visible, consisting of numerous small blue dots connected by thin white lines, forming a web-like structure. The overall aesthetic is technological and scientific.

Sports and Performance Science Portfolio

Christian Monteferrante, MSc., CSCS

- Project 1: Volleyball anual report – National University of La Matanza
 - Tools: Excell, Powerpoint, Jamovi (R)
 - Dataset: Real volleyball data
- Project 2: Edith Cowan University – Master of Exercise Science
 - Tools: Excel, Visual Basic
 - Dataset: ECU – Monitoring unit assignment
- Project 3: Big Data – Data Science course
 - Tools: Google Sheets, Google Looker Studio, Python, Deep Note
 - Datasets: “Services”, “Internet”, “Exams”

Project 1: University of La Matanza 2023

Anual Report



Volleyball D1 – Female

Christian Monteferrante, MSc., CSCS
Universidad Nacional de La Matanza



Personal Report



Excel



Powerpoint

Personal Report

1RM Estimated from Multiple Max Repetitions (xRM) - April to November 2023

Player

Assistance
92%

Bodyweight

66.7 kg ↓ -0.80

Absolute Values for Estimated 1RM

Back Squat
82.7 Kg ↓ -2.9

Deadlift
103.6 Kg ↑ 7.9

Lat Pulldown
66.7 Kg ↓ -3.1

Press
34.4 Kg ↑ 1.8

Pull Ups: 0.00

Relative to Bodyweight Values

The relative index refers to the relation between bodyweight (bw) and weight lifted. An index of 1 indicates 100% bw, 1.5, 150%bw

If any value is equal to 0, it means that the exercise was not tested

The last test was performed at the end of the Metropolitan Tournament and it corresponded to a maintenance phase

Very Good

Good

Acceptable

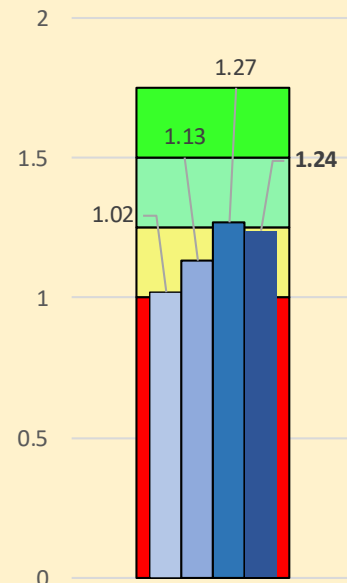
Poor

June

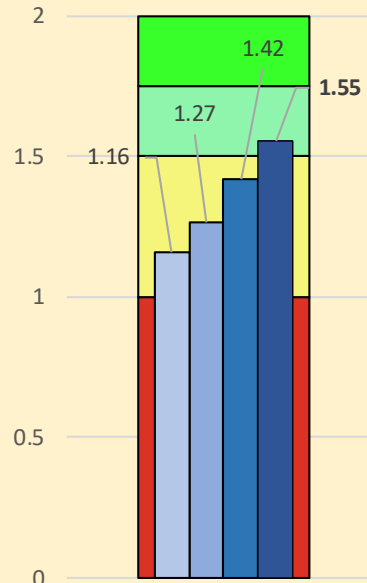
August

November

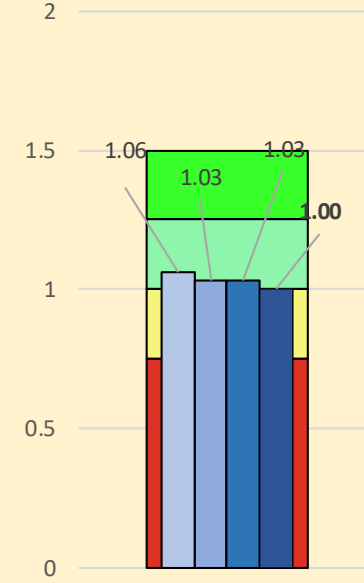
Back Squat



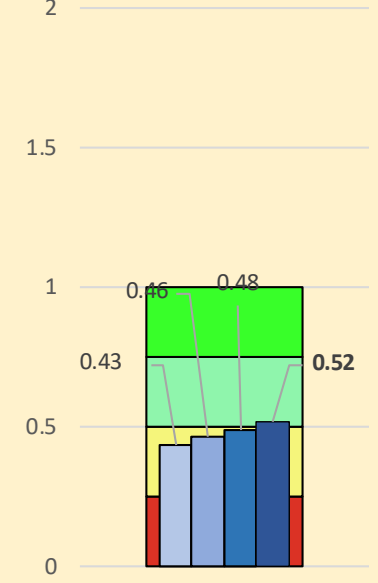
Deadlift



Lat Pulldown



Press



Backroom

Noviembre 2023

Fuerza												
Sent			Tiron		Despegue		Press					
Atleta	PC kg	S 1RM	S rel	TP 1RM	TP rel	D 1RM	D rel	P 1RM	P rel	Dominada	Asistencia	
C	60.2				0.00				0.00	no	77%	
C	66.7	82.7	1.24	66.7	1.00	103.6	1.55	34.4	0.52	si	92%	
D	69.9	88.0	1.26	76	1.09	111.6	1.60	36.4	0.52	si	99%	
G	72	91.9	1.28	55.6	0.77	90	1.25	34.0	0.47	no	76%	
L	61.3		0.00		0.00		0.00		0.00		78%	
L	64.4	81.5	1.27	74.9	1.16	87.9	1.36	36.1	0.56	si	91%	
N	59.1	71.0	1.20	59.8	1.01	88	1.49	31.3	0.53	no	98%	
R	61.3	83.3	1.36	64.5	1.05	92.4	1.51	30.1	0.49	si	86%	
S	61.2	100.6	1.64	79.8	1.30	98.3	1.61	40.0	0.65	si	97%	
S	89	103.5	1.16	73.7	0.83	151.8	1.71	39.2	0.44	no	88%	
V	79.5	104.4	1.31	81.4	1.02	132.7	1.67	40.0	0.50	no	94%	
F	72.5	98.9	1.36	84.3	1.16	111.1	1.53	40.0	0.55	si	87%	
V	63.9	88.4	1.38	63.9	1.00	83.3	1.30	34.9	0.55	si		
N	79.5	103.3	1.30	77.8	0.98	105.3	1.32	33.3	0.42	no	90%	

Personal Report - Anthropometrics

Athlete:

muscle-bone index

Ideally, we are looking for the green zone and as far as the red zone as possible. The cut-off point for female university volleyball players is 3.8 (arbitrary units)

sum of 6 skinfolds

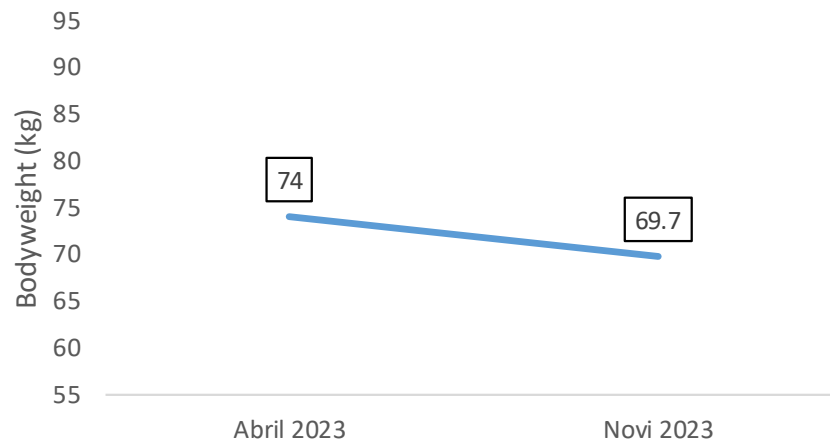
In this case, we aim to be below the 85 millimeters mark, indicated by the green zone

Important!!!

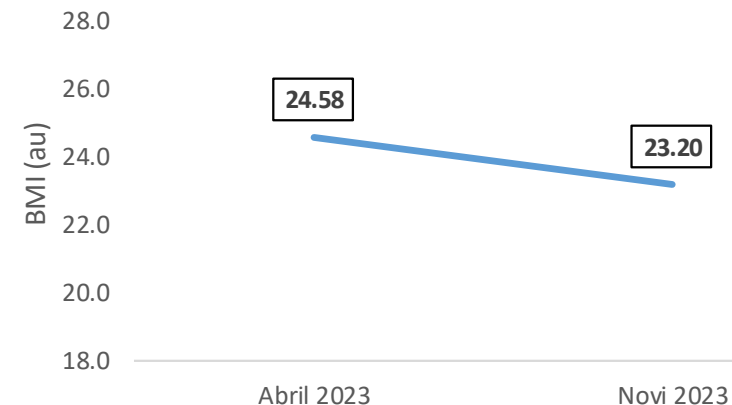
None of the markers indicated on the graphs is a predictor of performance

Anthropometric Data Evolution during 2023

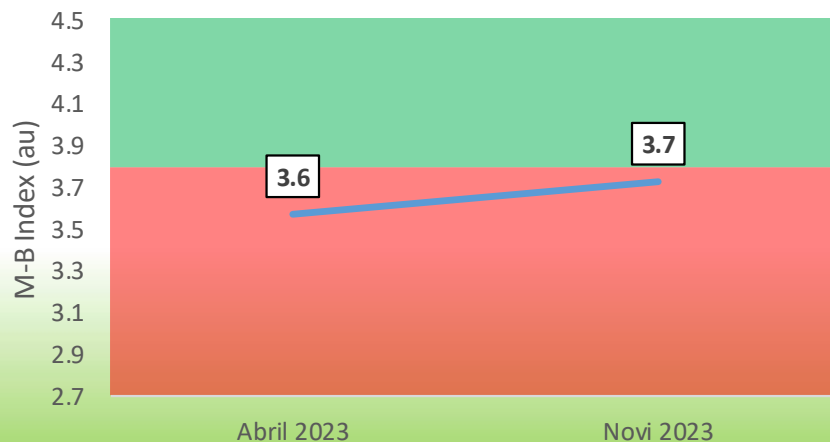
Bodyweight Evolution



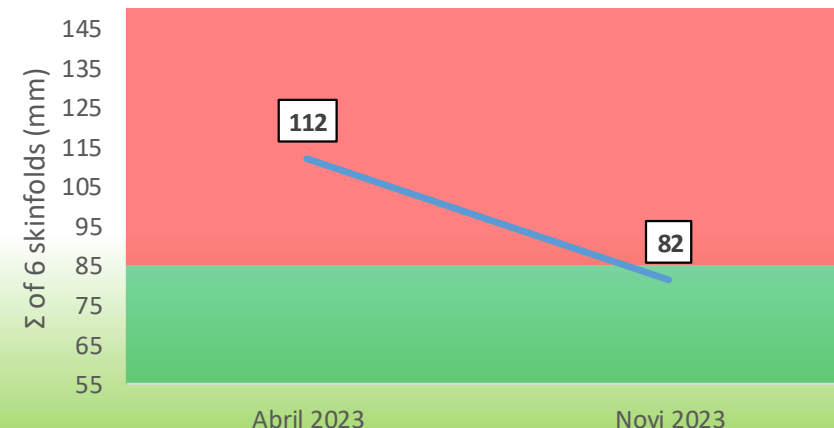
BMI Evolution



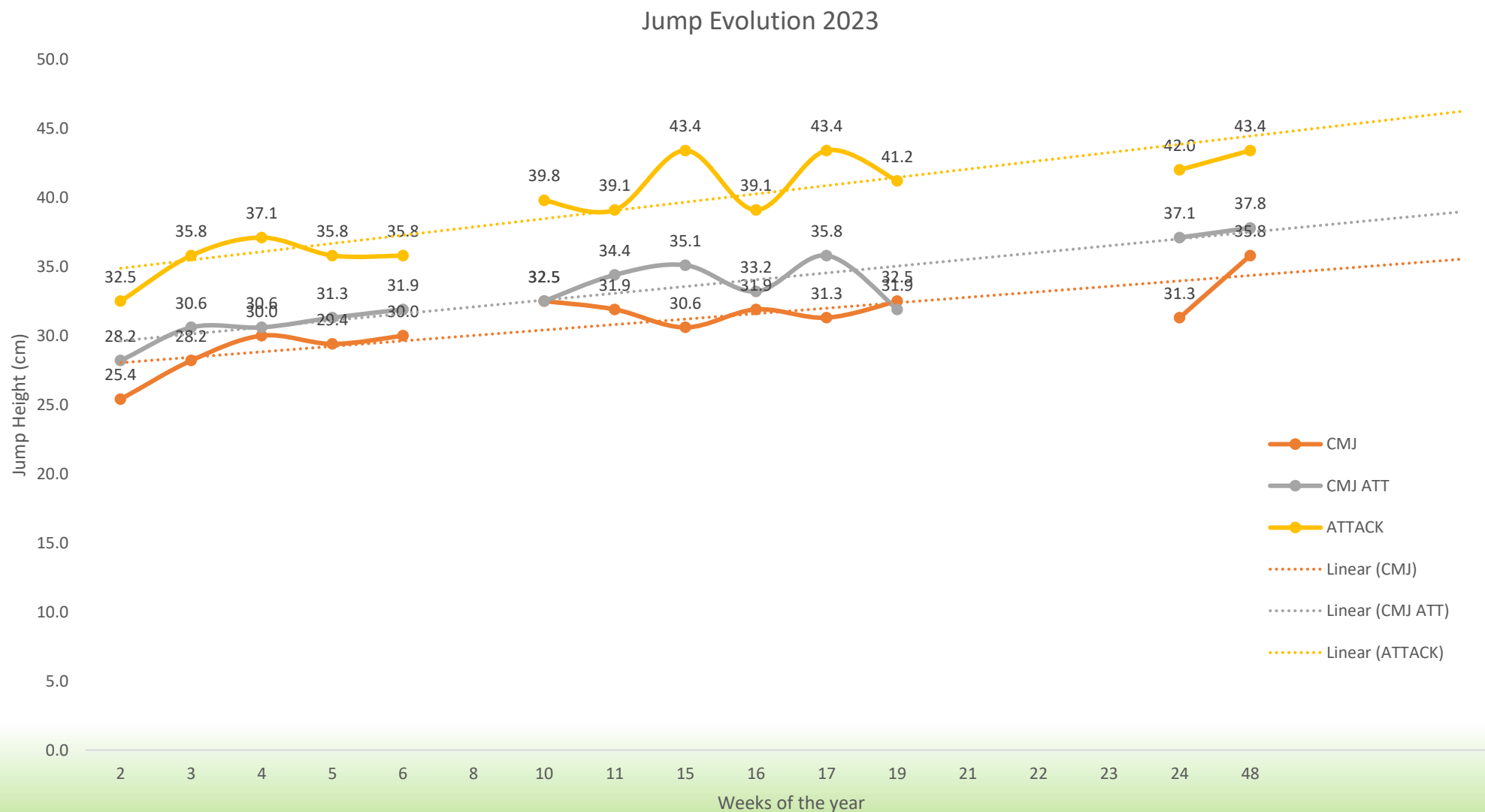
Muscle-Bone Index Evolution



Sum of 6 Skinfolds Evolution



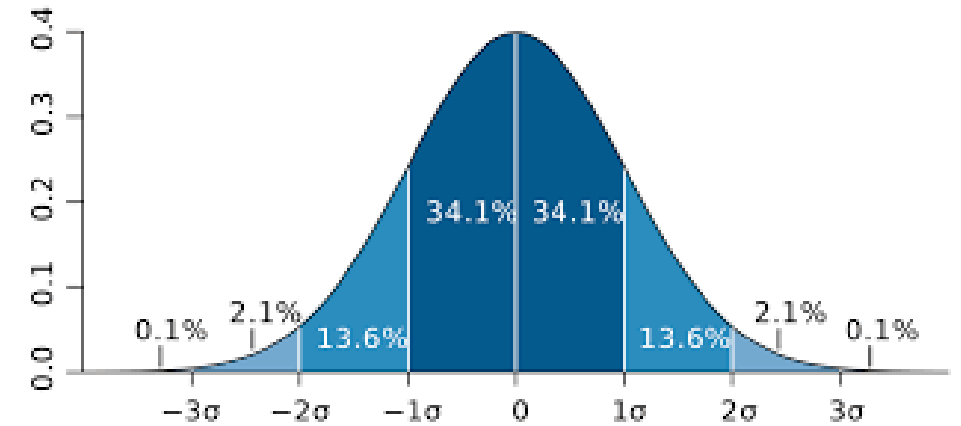
Personal Report - Jumps



Z-Scores

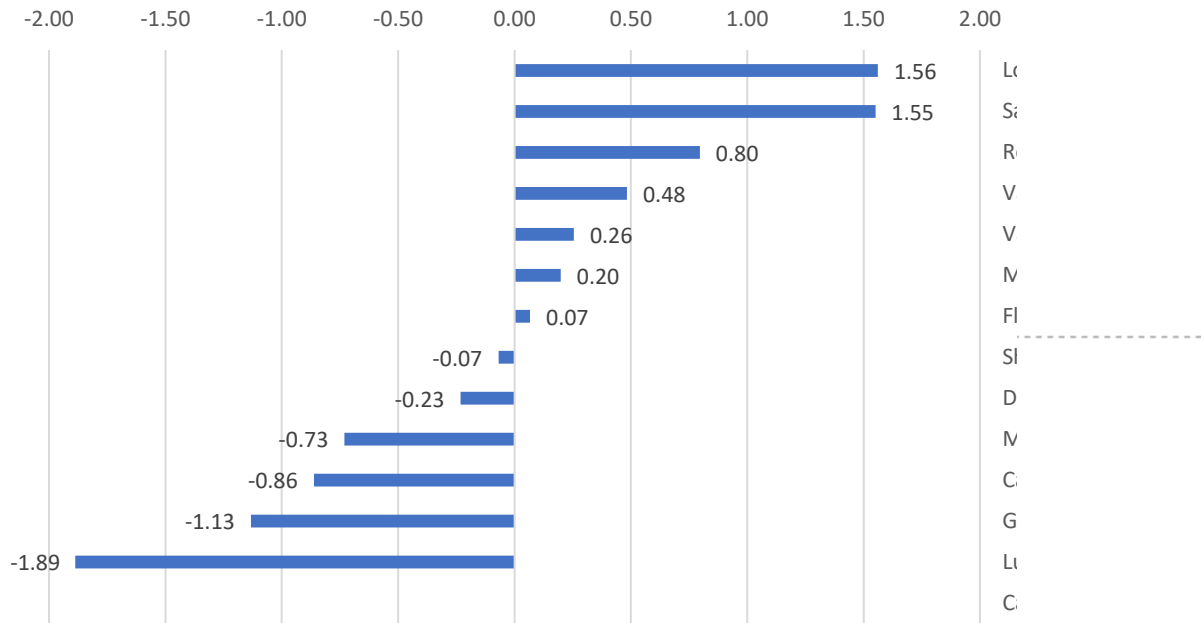
Z-score is a statistical measurement that describes a value's relationship to the mean of a group of values. Z-score is measured in terms of standard deviations from the mean. If a Z-score is 0, it indicates that the data point's score is identical to the mean score.

It is important for us so we can understand where our players are situated in relation to the group. On the following cases, the relative-to-bodyweight values were compared for the selected exercises.

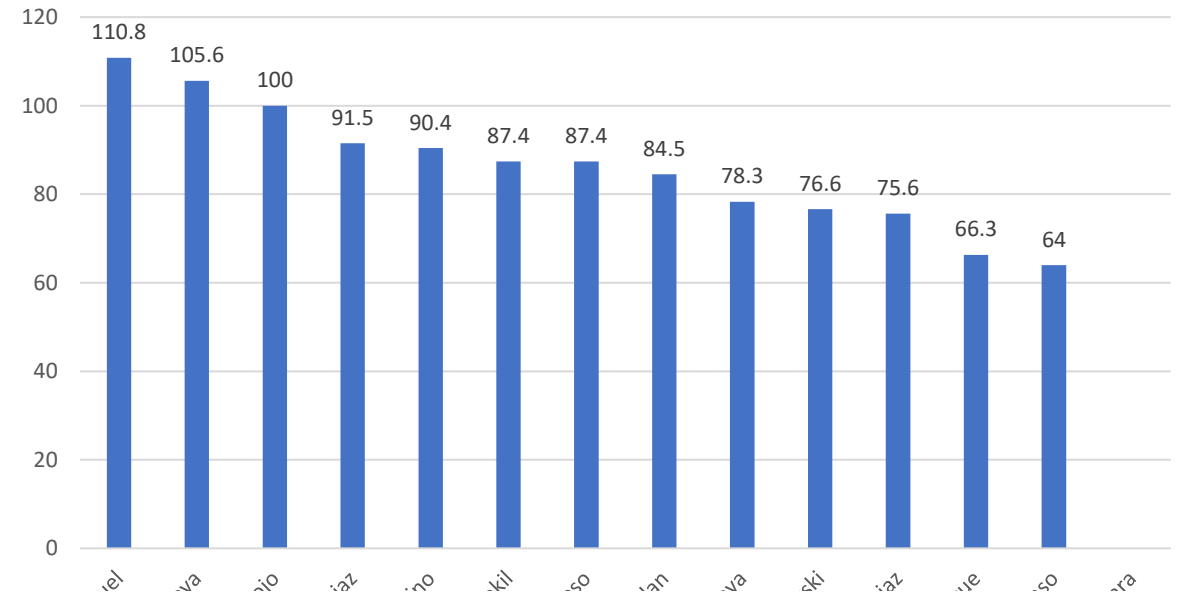


Squat

Relative Z-Score - Squat



Estimated 1RM - Squat



Backroom

Peso Corporal				Sentadillas				Tiron Palmar				Despegue				Press			
Testeo	Atleta	PC		Testeo	Atleta	TP	S rel	Testeo	Atleta	TP	TP rel	Testeo	Atleta	D	D rel	Testeo	Atleta	P IRl	P re
1ro - Abril	C	55.00		1ro - Abril	C			1ro - Abril	C	53.30	1.06	1ro - Abril	C			1ro - Abril	C	26.00	0.47
1ro - Abril	C	69.20		1ro - Abril	C	70.70	1.23	1ro - Abril	C	73.30	0.99	1ro - Abril	C	80.30	0.00	1ro - Abril	C	29.80	0.43
1ro - Abril	C	74.00		1ro - Abril	D	90.70	1.15	1ro - Abril	C	73.30	0.76	1ro - Abril	D			1ro - Abril	D	30.50	0.41
1ro - Abril	C	68.50		1ro - Abril	C	78.80	1.35	1ro - Abril	C	52.30	1.07	1ro - Abril	C			1ro - Abril	G	29.70	0.43
1ro - Abril	L	59.40		1ro - Abril	L	80.00	1.02	1ro - Abril	L	63.30	0.93	1ro - Abril	L	108.00	1.37	1ro - Abril	L	30.40	0.51
1ro - Abril	L	64.20		1ro - Abril	L	65.20	0.00	1ro - Abril	L	60.00	0.95	1ro - Abril	L	88.00	1.13	1ro - Abril	L	27.30	0.43
1ro - Abril	M	56.90		1ro - Abril	M			1ro - Abril	M	54.20	0.85	1ro - Abril	M	64.40	1.37	1ro - Abril	M	25.60	0.45
1ro - Abril	F	61.40		1ro - Abril	F	69.80	1.42	1ro - Abril	F	52.00	1.26	1ro - Abril	F	84.40	1.50	1ro - Abril	R	25.30	0.41
1ro - Abril	S	63.70		1ro - Abril	S	90.40	0.78	1ro - Abril	S	80.00	0.65	1ro - Abril	S	95.60	0.00	1ro - Abril	S	34.00	0.53
1ro - Abril	S	100.00		1ro - Abril	S	77.90	1.21	1ro - Abril	S	65.20	0.81	1ro - Abril	S			1ro - Abril	S		
1ro - Abril	S	88.20		1ro - Abril	S	106.60	1.09	1ro - Abril	S	71.40	1.01	1ro - Abril	S	115.20	1.20	1ro - Abril	S		
1ro - Abril	F	71.50		1ro - Abril	F	77.90	1.29	1ro - Abril	F	72.50	1.07	1ro - Abril	F	85.90	1.06	1ro - Abril	FI	36.40	0.51
1ro - Abril	V	60.90		1ro - Abril	V	78.30	1.18	1ro - Abril	V	65.20	1.06	1ro - Abril	V	64.40	1.52	1ro - Abril	V	30.40	0.50
1ro - Abril	V	77.80		1ro - Abril	V	91.80	0.00	1ro - Abril	V	82.30	0.00	1ro - Abril	V	118.20	0.00	1ro - Abril	V	35.90	0.46
1ro - Abril	M	76.20		1ro - Abril	M			1ro - Abril	M			1ro - Abril	M			1ro - Abril	M		
2do - Junio	C	58.00		2do - Junio	C		1.13	2do - Junio	C	56.30	1.03	2do - Junio	C		1.27	2do - Junio	C	28.00	0.48
2do - Junio	C	67.50		2do - Junio	C	76.60	1.22	2do - Junio	C	69.80	1.01	2do - Junio	C	85.50	1.35	2do - Junio	C	31.30	0.46
2do - Junio	C	71.50		2do - Junio	D	87.40	1.10	2do - Junio	C	72.30	0.81	2do - Junio	D	96.60	0.00	2do - Junio	D	32.90	0.46
2do - Junio	C	68.90		2do - Junio	C	75.60	1.47	2do - Junio	C	55.60	1.14	2do - Junio	C		1.82	2do - Junio	G	30.20	0.44
2do - Junio	L	59.40		2do - Junio	L	87.40	0.99	2do - Junio	L	67.90	1.12	2do - Junio	L	108.00	1.42	2do - Junio	L	34.70	0.58
2do - Junio	L	64.50		2do - Junio	L	64.00	1.15	2do - Junio	L	72.50	0.94	2do - Junio	L	91.80	1.35	2do - Junio	L	30.00	0.47
2do - Junio	M	57.50		2do - Junio	M	66.30	1.37	2do - Junio	M	54.20	1.02	2do - Junio	M	77.40	1.51	2do - Junio	M	25.90	0.45
2do - Junio	F	61.90		2do - Junio	F	84.50	1.47	2do - Junio	F	63.40	1.47	2do - Junio	F	93.30	1.57	2do - Junio	R	26.60	0.43
2do - Junio	S	61.50		2do - Junio	S	90.40	1.24	2do - Junio	S	90.40	0.78	2do - Junio	S	96.40	1.50	2do - Junio	S	38.00	0.62
2do - Junio	S	89.00		2do - Junio	S	110.80	1.32	2do - Junio	S	69.80	1.06	2do - Junio	S	133.30	1.81	2do - Junio	S	32.90	0.37
2do - Junio	V	79.90		2do - Junio	V	105.60	1.26	2do - Junio	V	84.40	1.20	2do - Junio	V	144.90	1.51	2do - Junio	V	38.30	0.48
2do - Junio	F	72.40		2do - Junio	F	91.50	1.29	2do - Junio	F	86.90	1.01	2do - Junio	F	109.30	1.19	2do - Junio	FI	38.30	0.53
2do - Junio	V	60.70		2do - Junio	V	78.30	1.28	2do - Junio	V	61.10	0.89	2do - Junio	V	72.30	0.00	2do - Junio	V	31.30	0.52
2do - Junio	M	78.00		2do - Junio	M	100.00	0.00	2do - Junio	M	69.80	0.97	2do - Junio	M		0.00	2do - Junio	M	27.80	0.36
3ro - Agosto	C	58.00		3ro - Agosto	C		1.27	3ro - Agosto	C	56.30	1.03	3ro - Agosto	C		1.42	3ro - Agosto	C	28.00	0.48
3ro - Agosto	C	67.50		3ro - Agosto	C	85.60	1.26	3ro - Agosto	C	69.80	1.05	3ro - Agosto	C	95.70	1.63	3ro - Agosto	C	32.60	0.48
3ro - Agosto	C	71.50		3ro - Agosto	D	90.00	1.35	3ro - Agosto	C	75.30	0.81	3ro - Agosto	D	116.90	1.22	3ro - Agosto	D	36.10	0.50
3ro - Agosto	C	68.90		3ro - Agosto	C	93.30	2.16	3ro - Agosto	C	55.60	1.14	3ro - Agosto	C	84.40	1.94	3ro - Agosto	G	34.40	0.50
3ro - Agosto	L	59.40		3ro - Agosto	L	128.10	0.99	3ro - Agosto	L	67.90	1.12	3ro - Agosto	L	115.20	1.42	3ro - Agosto	L	34.70	0.58
3ro - Agosto	L	64.50		3ro - Agosto	L	64.00	1.34	3ro - Agosto	L	72.50	0.94	3ro - Agosto	L	91.80	1.54	3ro - Agosto	L	30.00	0.47
3ro - Agosto	M	57.50		3ro - Agosto	M	77.20	1.37	3ro - Agosto	M	54.20	1.02	3ro - Agosto	M	88.60	1.51	3ro - Agosto	M	25.90	0.45
3ro - Agosto	F	61.90		3ro - Agosto	F	84.50	1.47	3ro - Agosto	F	63.40	1.47	3ro - Agosto	F	93.30	1.57	3ro - Agosto	R	26.60	0.43
3ro - Agosto	S	61.50		3ro - Agosto	S	90.40	1.29	3ro - Agosto	S	90.40	0.78	3ro - Agosto	S	96.40	1.60	3ro - Agosto	S	38.00	0.62
3ro - Agosto	S	89.00		3ro - Agosto	S	115.00	1.32	3ro - Agosto	S	69.80	1.06	3ro - Agosto	S	142.50	1.81	3ro - Agosto	S	37.40	0.42
3ro - Agosto	V	79.90		3ro - Agosto	V	105.60	1.37	3ro - Agosto	V	84.40	1.16	3ro - Agosto	V	144.90	1.52	3ro - Agosto	V	38.30	0.48
3ro - Agosto	F	72.40		3ro - Agosto	F	98.90	1.44	3ro - Agosto	F	84.20	1.05	3ro - Agosto	F	109.80	1.39	3ro - Agosto	FI	43.40	0.60
3ro - Agosto	V	60.70		3ro - Agosto	V	87.20	1.48	3ro - Agosto	V	63.90	0.89	3ro - Agosto	V	84.40	1.42	3ro - Agosto	V	34.40	0.57
3ro - Agosto	M	78.00		3ro - Agosto	M	115.20	0.00	3ro - Agosto	M	69.80	0.00	3ro - Agosto	M	110.50	0.00	3ro - Agosto	M	31.30	0.40
4to - Noviembre	C	60.20		4to - Noviembre	C		1.24	4to - Noviembre	C		1.00	4to - Noviembre	C		1.55	4to - Noviembre	C		0.00
4to - Noviembre	C	66.70		4to - Noviembre	C	82.70	1.26	4to - Noviembre	C	66.70	1.09	4to - Noviembre	C	103.60	1.60	4to - Noviembre	C	34.40	0.52
4to - Noviembre	C	69.90		4to - Noviembre	D	88.00	1.33	4to - Noviembre	C	76.00	0.81	4to - Noviembre	D	111.60	1.31	4to - Noviembre	D	36.40	0.52
4to - Noviembre	C	68.90		4to - Noviembre	C	91.90	0.00	4to - Noviembre	C	55.60	0.00	4to - Noviembre	C	90.00	0.00	4to - Noviembre	G	34.00	0.49
4to - Noviembre	L	61.30		4to - Noviembre	L		1.27	4to - Noviembre	L		1.16	4to - Noviembre	L		1.36	4to - Noviembre	L		0.00
4to - Noviembre	L	64.40		4to - Noviembre	L	81.50	1.20	4to - Noviembre	L	74.90	1.01	4to - Noviembre	L	87.90	1.49	4to - Noviembre	L	36.10	0.56
4to - Noviembre	M	59.10		4to - Noviembre	M	71.00	1.36	4to - Noviembre	M	59.80	1.05	4to - Noviembre	M	88.00	1.51	4to - Noviembre	M	31.30	0.53
4to - Noviembre	F	61.30		4to - Noviembre	F	83.30	1.64	4to - Noviembre	F	64.50	1.30	4to - Noviembre	F	92.40	1.61	4to - Noviembre	R	30.10	0.49
4to - Noviembre	S	61.20		4to - Noviembre	S	100.60	1.16	4to - Noviembre	S	79.80	0.83	4to - Noviembre	S	98.30	1.71	4to - Noviembre	S	40.00	0.65
4to - Noviembre	S	89.00		4to - Noviembre	S	103.50	1.31	4to - Noviembre	S	73.70	1.02	4to - Noviembre	S	151.80	1.67	4to - Noviembre	S	39.20	0.44
4to - Noviembre	V	79.50		4to - Noviembre	V	104.40	1.36	4to - Noviembre	V		1.16	4to - Noviembre	V	132.70	1.53	4to - Noviembre	V	40.00	0.50
4to - Noviembre	F	72.50		4to - Noviembre	F	98.90	1.38	4to - Noviembre	F	84.30	1.00	4to - Noviembre	F	111.10	1.30	4to - Noviembre	FI	40.00	0.55

Relative Squat

L	1.47
S	1.47
R	1.37
V	1.32
V	1.29
N	1.28
F	1.26
S	1.24
C	1.22
N	1.15
C	1.13
G	1.10
L	0.99
C	N/A

Team Evolution

jamovi Stats.
Open.
Now.

 **R Programming**



Variables

Data

Analyses

Edit



Exploration



T-Tests



ANOVA



Regression



Frequencies



Factor



esci



Flexplot



Base R



R



Modules

	Equipo	Fecha	Atleta	PC kg	Sent 1RM	Sent rel	Tir
1	UNLaM	Abril 2023	C	55.0			
2	UNLaM	Abril 2023	C	69.2	70.7	1.02	
3	UNLaM	Abril 2023	D	74.0	90.7	1.23	
4	UNLaM	Abril 2023	G	68.5	78.8	1.15	
5	UNLaM	Abril 2023	L	59.4	80.0	1.35	
6	UNLaM	Abril 2023	L	64.2	65.2	1.02	
7	UNLaM	Abril 2023	M	56.9			
8	UNLaM	Abril 2023	R	61.4	69.8	1.14	
9	UNLaM	Abril 2023	S	63.7	90.4	1.42	
10	UNLaM	Abril 2023	S	100.0	77.9	0.78	
11	UNLaM	Abril 2023	S	88.2	106.6	1.21	
12	UNLaM	Abril 2023	F	71.5	77.9	1.09	
13	UNLaM	Abril 2023	V	60.9	78.3	1.29	
14	UNLaM	Abril 2023	V	77.8	91.8	1.18	
15	UNLaM	Abril 2023	M	76.2			
16	UNLaM	Junio 2023	C	58.0			
17	UNLaM	Junio 2023	C	67.5	76.6	1.13	
18	UNLaM	Junio 2023	D	71.5	87.4	1.22	
19	UNLaM	Junio 2023	G	68.9	75.6	1.10	
20	UNLaM	Junio 2023	L	59.4			
21	UNLaM	Junio 2023	L	64.5			
22	UNLaM	Junio 2023	M	57.5	66.3	1.15	
23	UNLaM	Junio 2023	R	61.9	84.5	1.37	
24	UNLaM	Junio 2023	S	61.5	90.4	1.47	
25	UNLaM	Junio 2023	S	89.0	110.8	1.24	
26	UNLaM	Junio 2023	F	79.9	105.6	1.32	
27	UNLaM	Junio 2023	V	72.4	91.5	1.26	
28	UNLaM	Junio 2023	V	60.7	78.3	1.29	

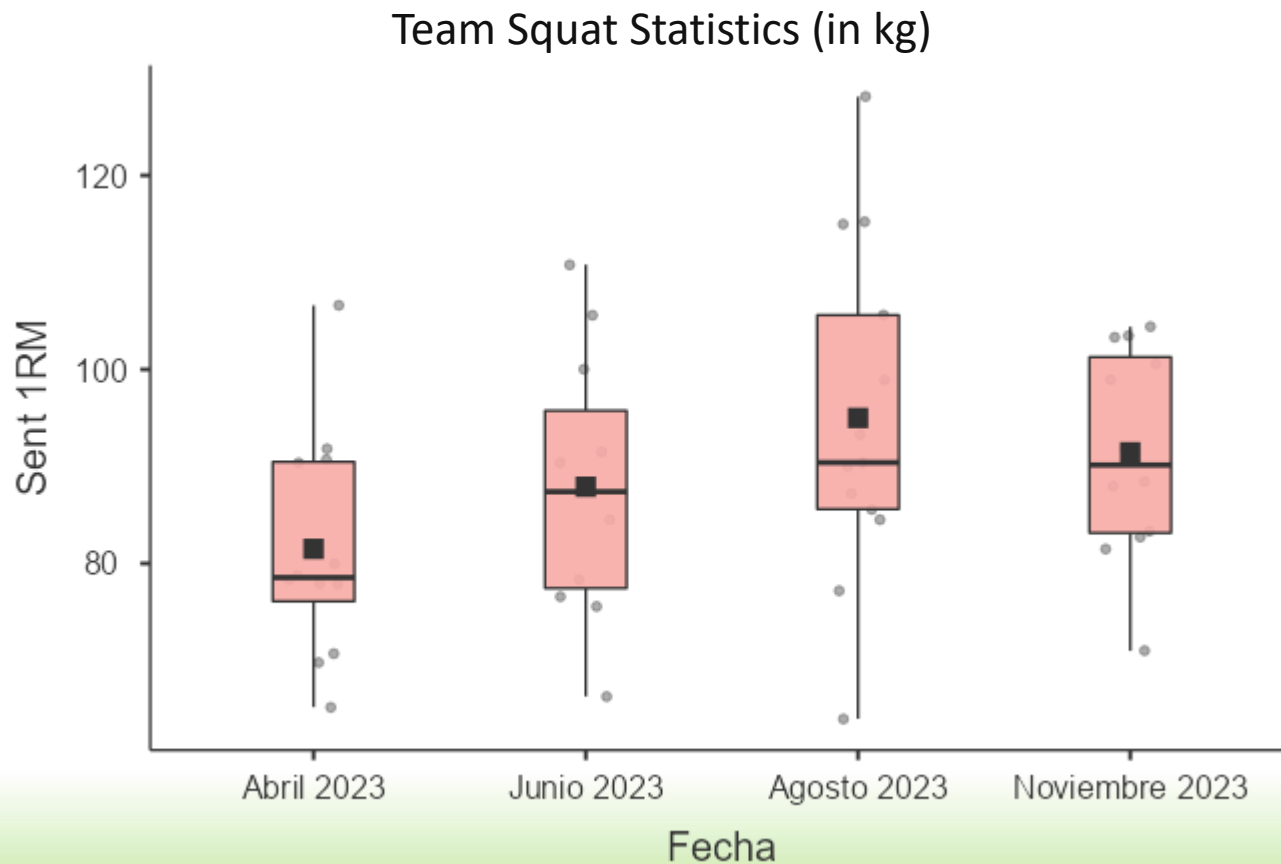
Results

Descriptives

Descriptives

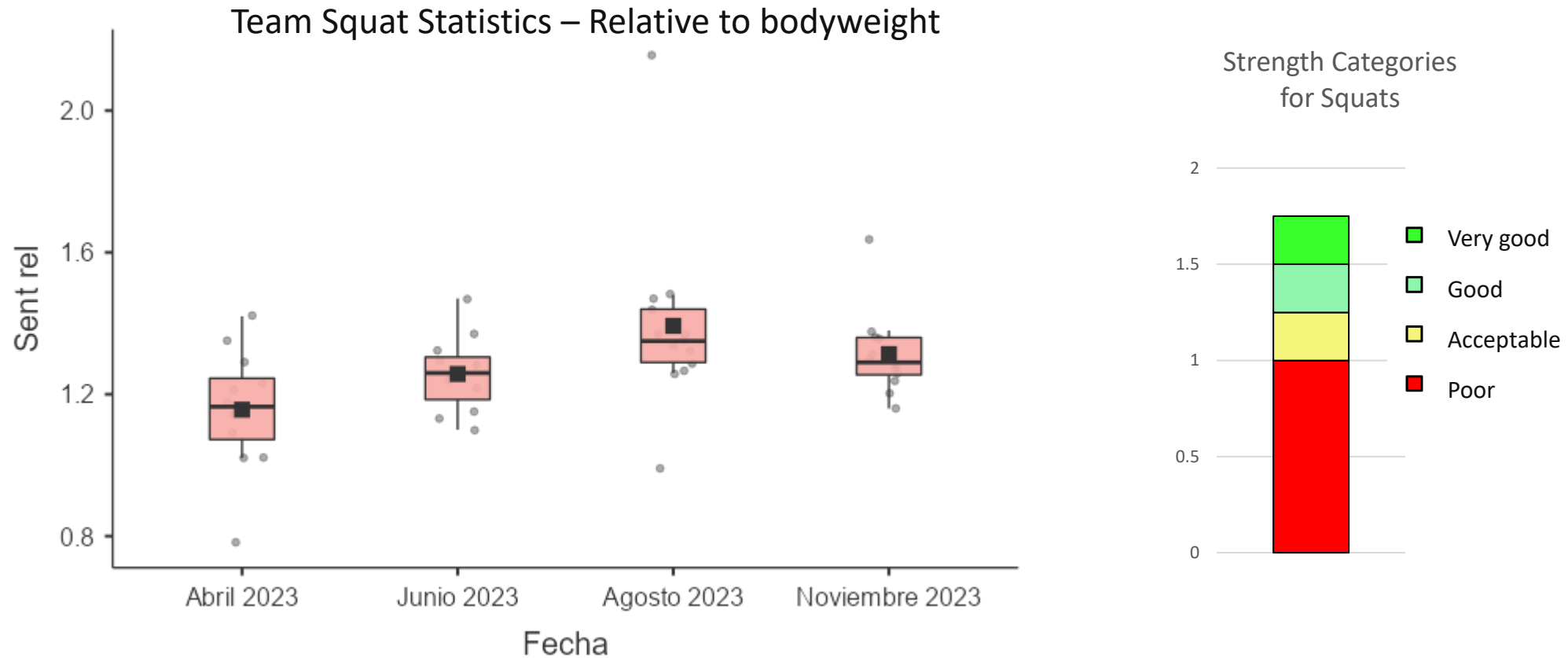
	Fecha	PC kg	Sent 1RM	Despegue 1RM	Press 1RM
N	Abril 2023	15	12	10	12
	Junio 2023	14	11	9	11
	Agosto 2023	14	13	13	14
	Noviembre 2023	14	12	12	12
Missing	Abril 2023	0	3	5	3
	Junio 2023	0	3	5	3
	Agosto 2023	0	1	1	0
	Noviembre 2023	0	2	2	2
Mean	Abril 2023	69.8	81.5	90.4	30.1
	Junio 2023	67.9	87.9	101	32.8
	Agosto 2023	67.9	95.0	106	33.6
	Noviembre 2023	68.6	91.5	105	35.8
Std. error mean	Abril 2023	3.15	3.33	6.00	1.09
	Junio 2023	2.51	4.10	8.13	1.31
	Agosto 2023	2.51	4.82	5.59	1.33
	Noviembre 2023	2.40	3.11	5.85	0.992
95% CI mean lower bound	Abril 2023	63.6	75.0	78.7	28.0
	Junio 2023	63.0	79.9	85.1	30.2
	Agosto 2023	63.0	85.6	94.8	31.1
	Noviembre 2023	63.9	85.4	93.2	33.9
95% CI mean upper bound	Abril 2023	76.0	88.0	102	32.2
	Junio 2023	72.8	96.0	117	35.3

Squat



The team's average squat strength increased throughout the year during the strength and power phases of the sport calendar. After the maintenance phase, which coincided with the Metropolitan Tournament, the strength level for this exercise slightly decreased within the expected parameters. However, the team finished the year with higher lower body strength compared to the beginning of the year. This places them in a more advantageous situation to tackle 2024 and the upcoming sport demands.

Squat



The relative strength for the squat shows a similar trend compared to the absolute values. On average, the team is better equipped to start 2024 in the category considered 'Good', which is higher than 125% of their own bodyweight, compared to April 2023 when their values were 'Acceptable', or between 100% and 125% of their own weight.

Project 2: Edith Cowan University

Edith Cowan University

Monitoring Unit (Dr. Greg Haff) - 2021





Excel



Visual Basic

Home Screen – Quick accesses



Data Input

Training

Physical

Wellness

Match

Medical

Updates

Dashboards

Individuals

Weekly

Team

Print Reports

Physical

Training

Games

Wellness & Medical

Databases

Training

Physical

Wellness

Match

Medical

Updates

C-Panel

Data Input Tab: Training information

Daily Training Questionnaire

Date (dd-mm-yyyy)

Name

Kendall

Training Type

Endurance

Duration (minutes)

78

Session RPE

sRPE

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☒ 8 ☐ 9 ☐ 10

1 - "aint notin' but a peanut" ; 10 - "completely extenuating"

Training



Home



Upload

Controls for Dr. H:



Database

Hide Columns

Individual Dashboard

Controls

Select athlete to
visualize data:

French



Home

French



Medical

Illness Count	Injury Count	Other Medical Categories
0	0	0

Overall Availability Average

Player	3.0
Team	2.7

Weekly Average Availability



Performance Profile: French

Wellness

	Sleep Quality	Sleep Hours	Soreness	Stress	Fatigue	Motivation	Wellness Score
Average of the last 7 days	3.0	7.7	4.7	4.7	4.7	7.5	32.2
Average of the last 30 days	3.3	8.0	4.4	4.4	4.4	5.4	29.8
Overall player average	3.0	7.0	4.6	4.6	4.6	5.1	29.0

Training Load and Wellness



Training Duration



Rate of Perceived Exertion



Physical

	Fitness	Lower Body Strength	Power Clean	Upper Body Strength	Speed 10 mts	Jump Power
Rating	Good	Poor	Poor	Poor	Excellent	Ok
Result	2700	105	68.25	78.75	1.75	59
Band	4	2	2	2	5	4

Fitness



Lower Body Strength



Power Clean



Upper Body Strength



Speed 10 mts



Jump Power

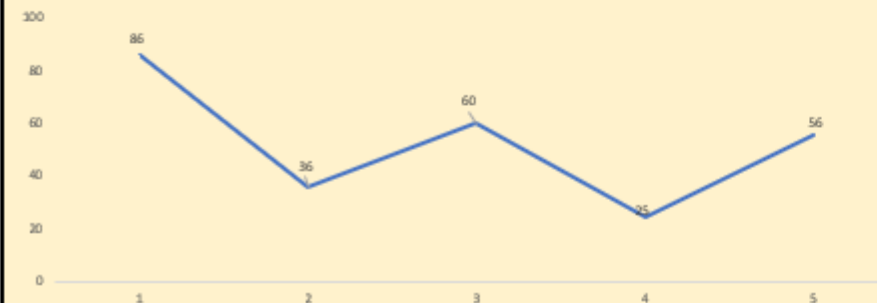


Updates

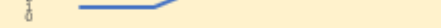
Date	Priority	Area	Comment
30/04/2022	2	Performance	Good COD technique
29/04/2022	2	Medical	Covid 19 positive
27/04/2022	2	Performance	good passes session
26/04/2022	2	Medical	Covid 19 positive
22/04/2022	2	Performance	Didn't hit tackle # goal

Match Performance

Match Performance Rating: French



Shots



Tackles



Passes



Receives



Errors




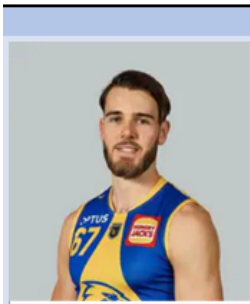
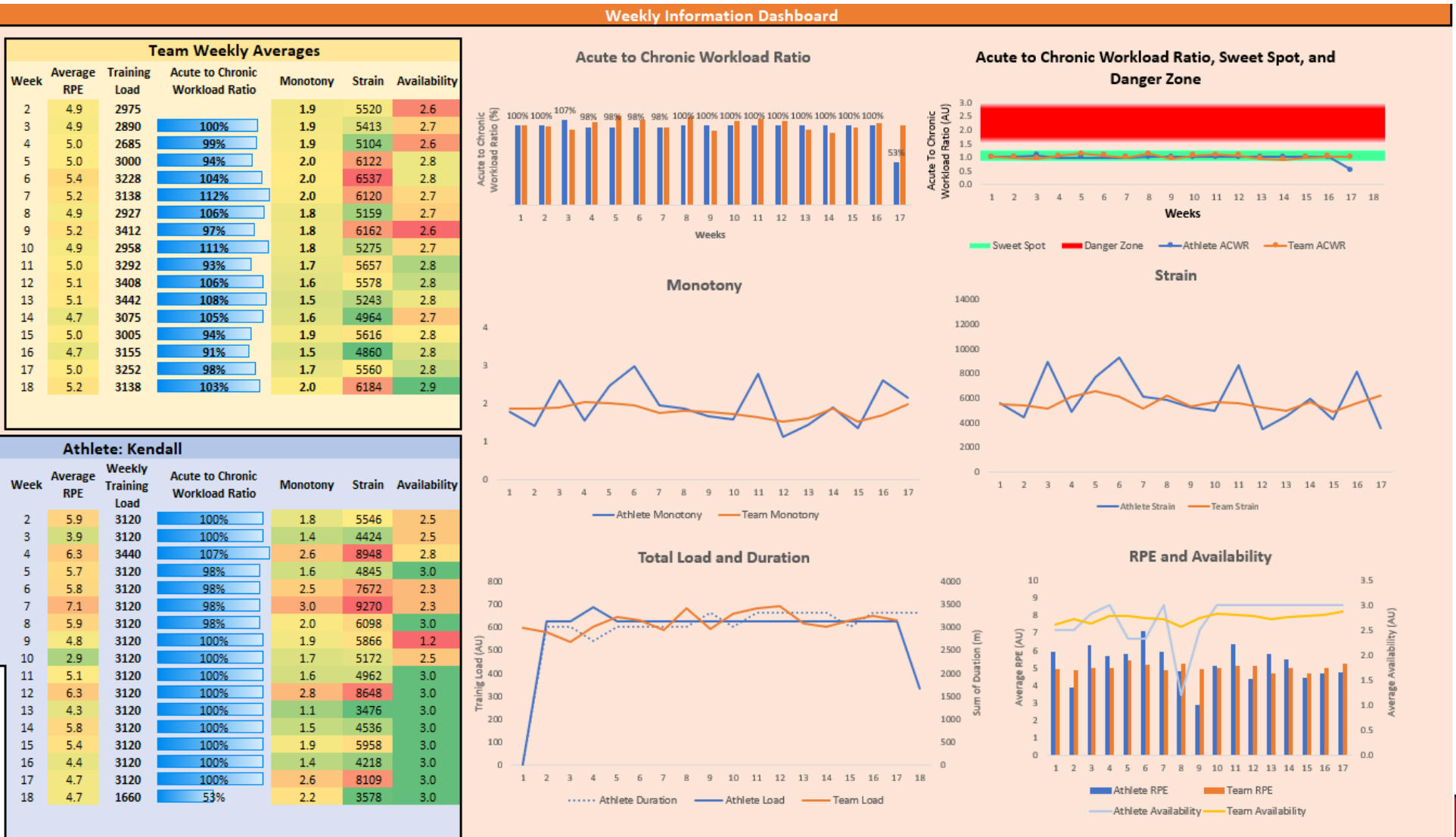
Weekly Dashboard

Controls

To plot "only team", leave blank. To plot "vs athlete", select:

Kendall

 Home



Team Dashboard

Controls



Home

Select a combination of weeks and days to visualize data:

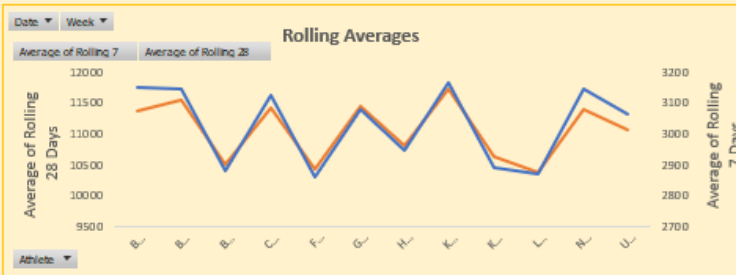
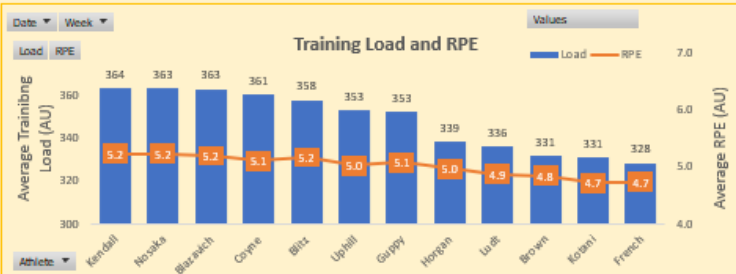
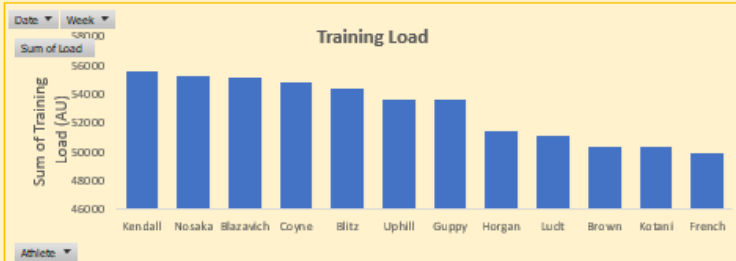
Week

2	3	4	5
6	7	8	9
10	11	12	13
14	15	16	17
18			

Date

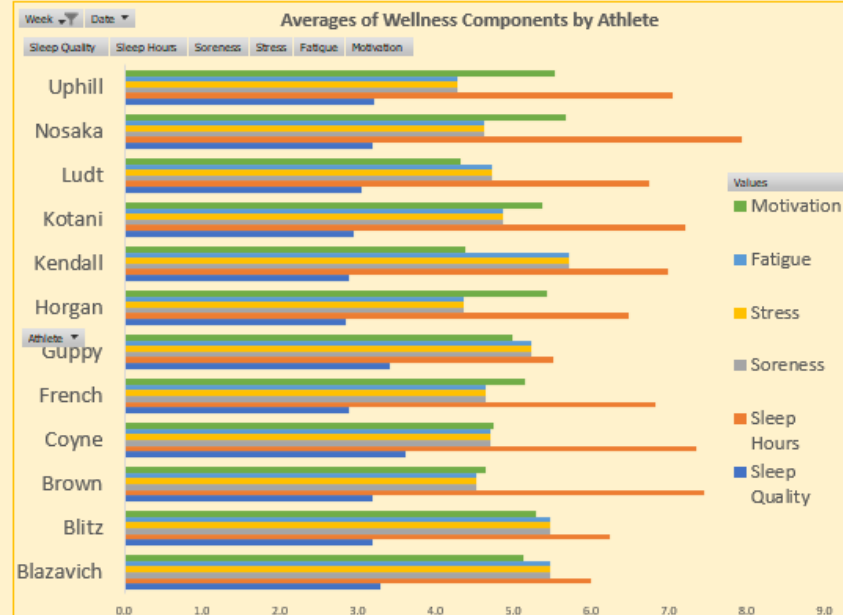
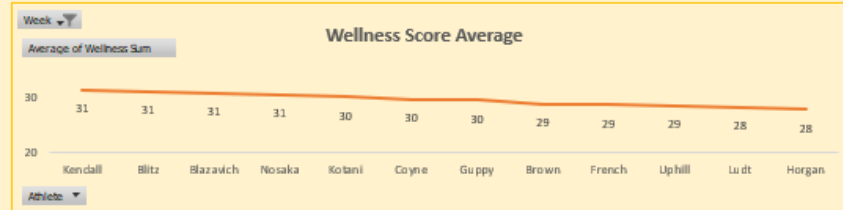
03-01-22	04-01-22	05-01-22
06-01-22	07-01-22	08-01-22
10-01-22	11-01-22	12-01-22
13-01-22	14-01-22	15-01-22
17-01-22	18-01-22	19-01-22
20-01-22	21-01-22	22-01-22
24-01-22	25-01-22	26-01-22
27-01-22	28-01-22	29-01-22

Load



Team Dashboard

Wellness



Physical Performance

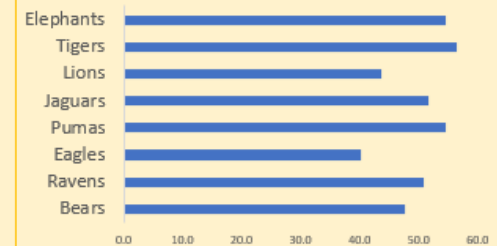
Latest Vs Overall Averages Latest Overall



Match Performance



Average Athlete Rating per Match



Wellness and Medical Report

Controls

Select athlete to visualize

Blitz

Select athlete and week to print

1

2

3

4

5

6

7


8


9

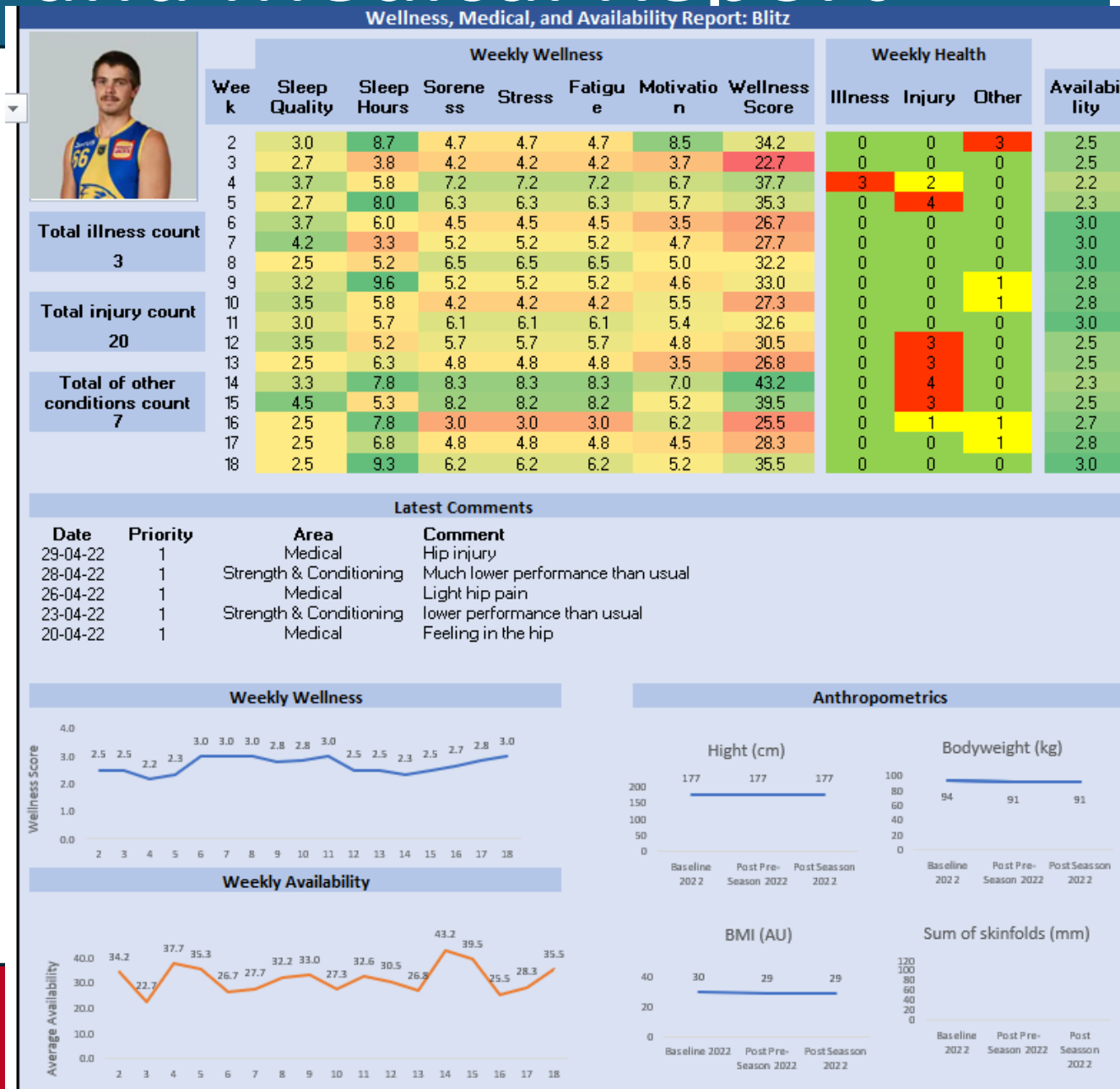
10

11

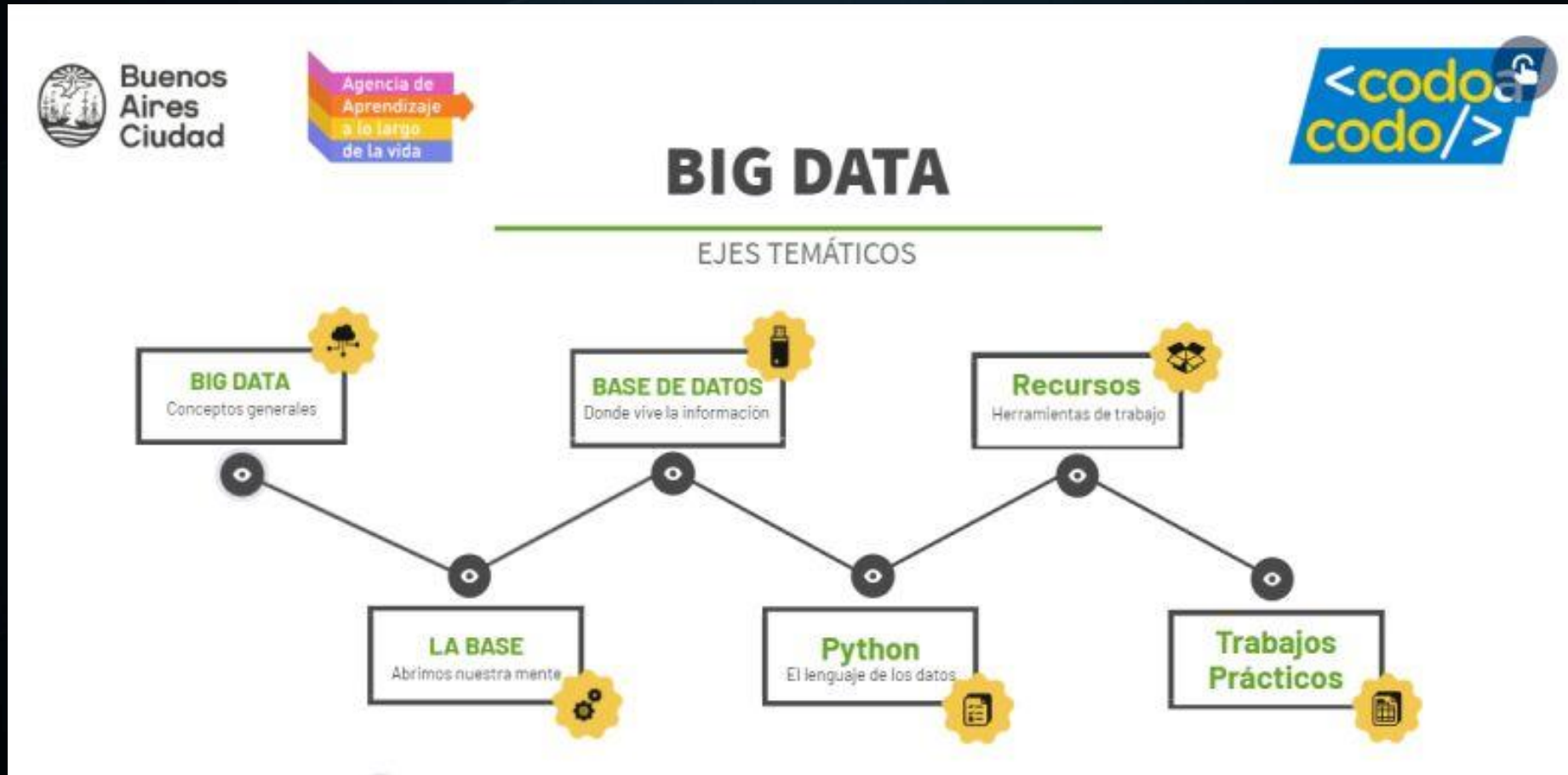
12

 **Print**

 **Home**



Project 3: Big Data – 2023



Sheets + Looker Studio

Databases



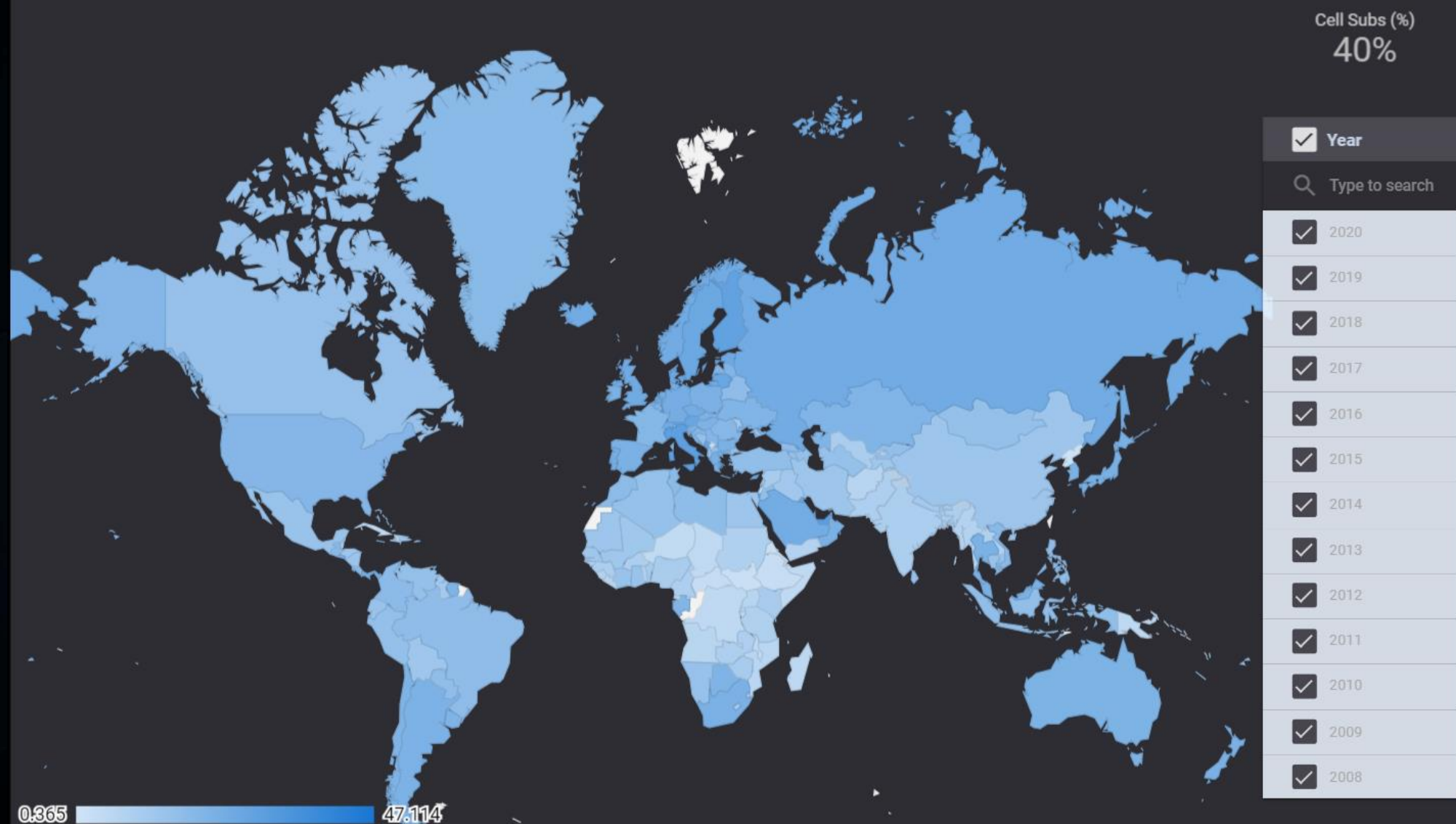
- "Services"
- "Internet"



Visualization



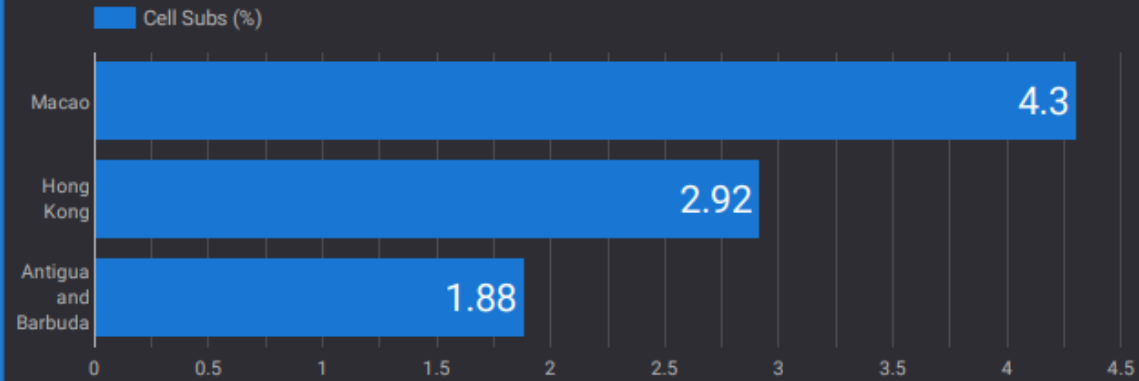
World mobile subscriptions percentage by country, from 1980 to 2020



Over 100% denotes more than one mobile subscription per person

Country with the most mobile lines per person in 2020: Macao

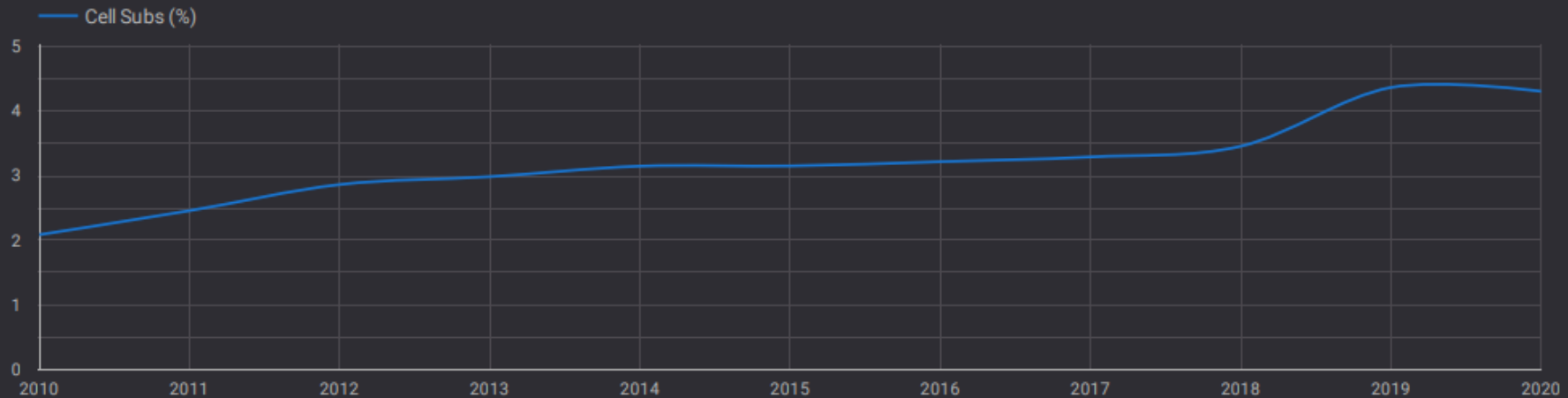
Top three countries with the most mobile lines per person



Macao's average mobile line's ownership per person:

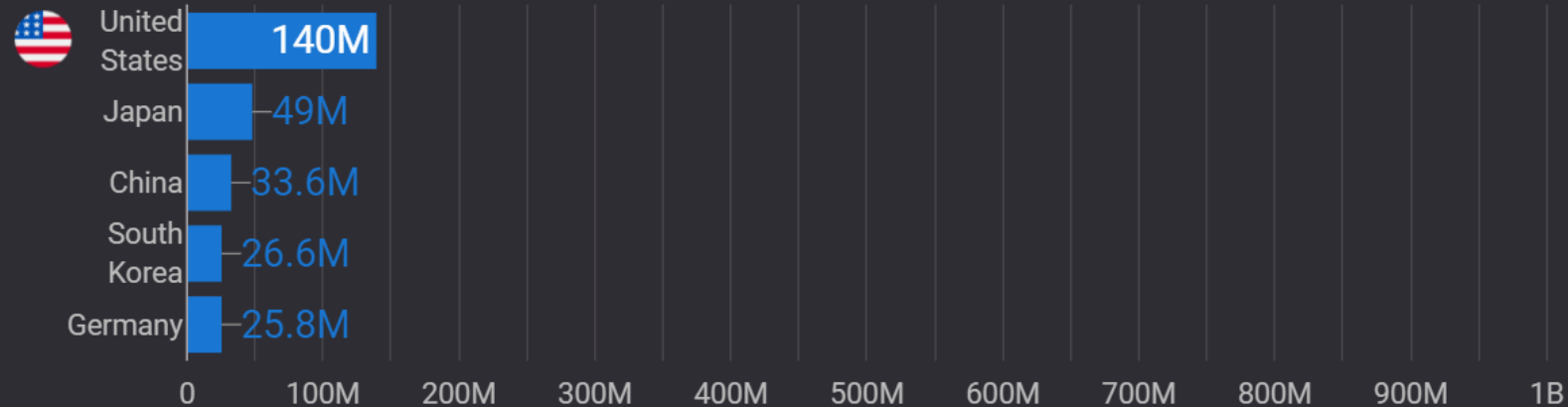
4.3

Macao's cell-phones ownership for 2010-2020

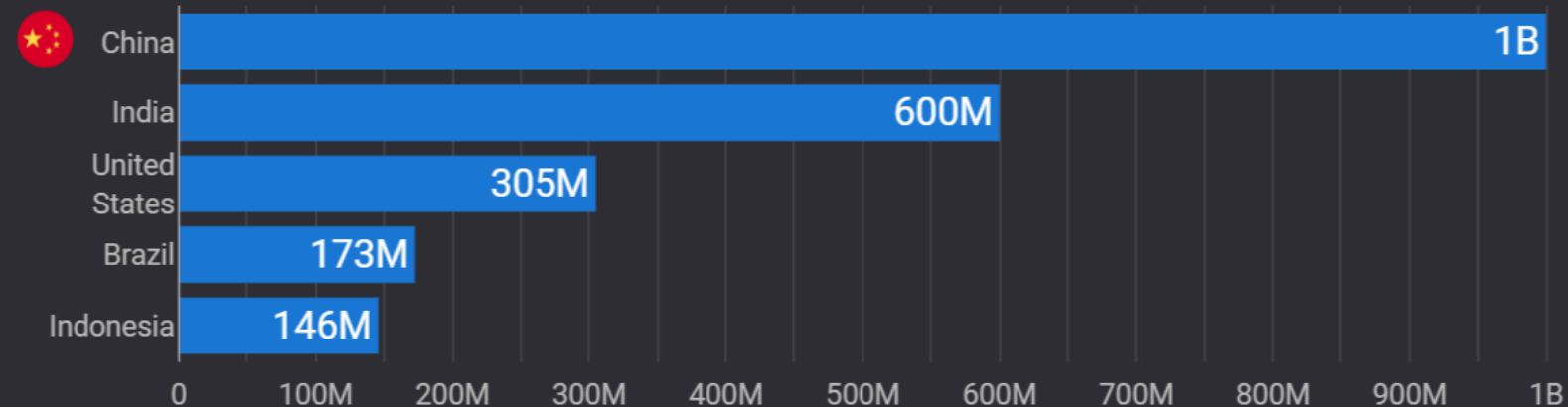


Top five countries with the highest number of internet users

2001



2020





Deepnote



python™

Dataset: “exams”

Basic Operations

Review data types

```
1 df.dtypes
```

id	object
gender	object
race/ethnicity	object
parental level of education	object
lunch	object
employed	object
test preparation course	object
math score	float64
physics score	float64
chemistry score	float64
algebra_score	float64
dtype:	object

Drop duplicates

```
::
```

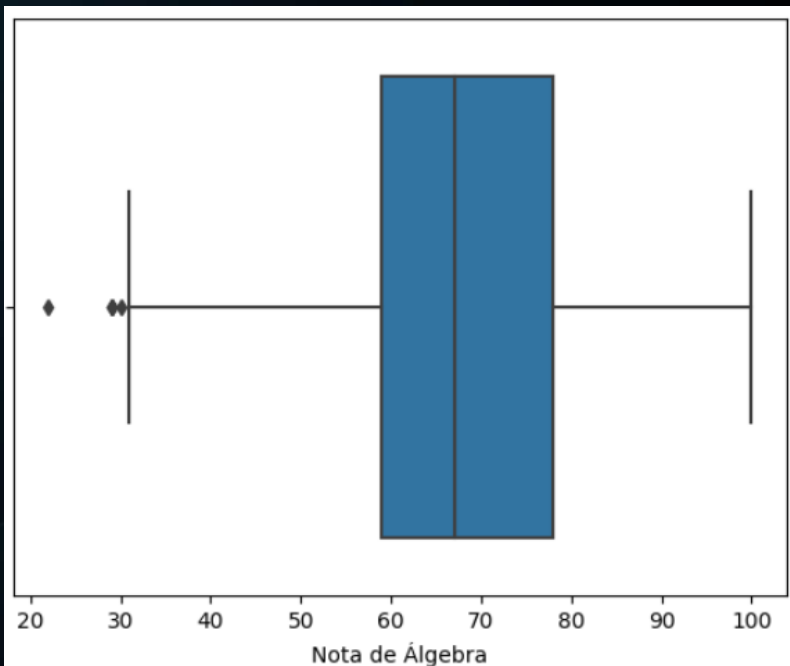
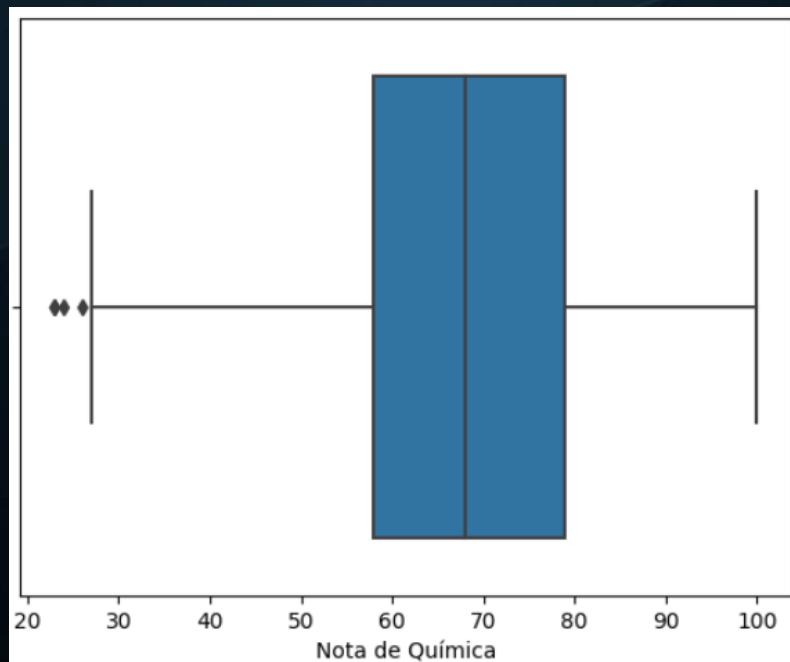
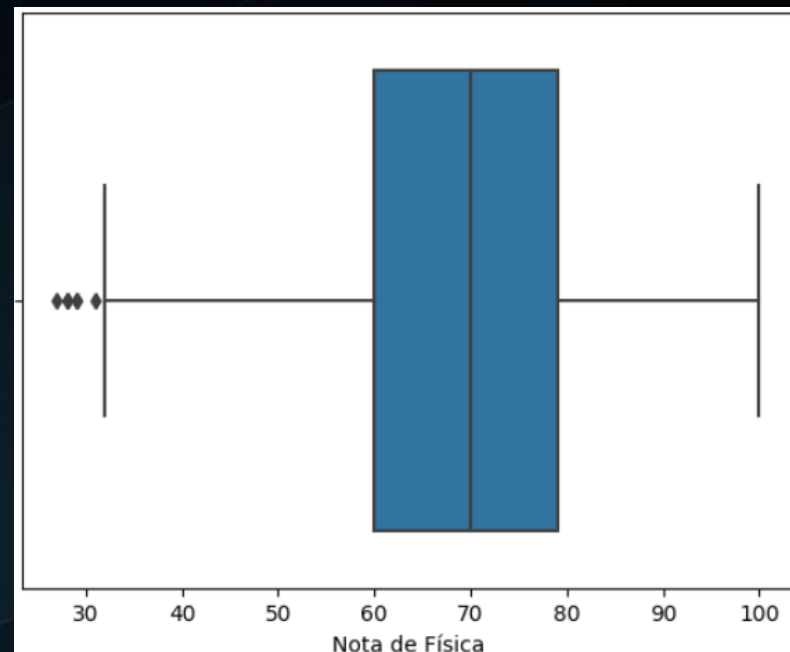
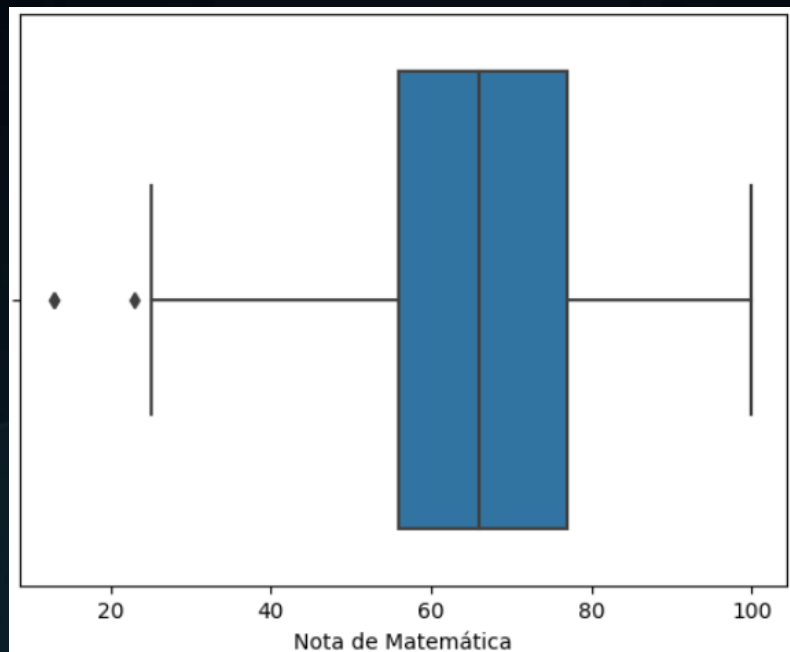
```
1 print(f'Original: {df.id.count()} filas')
2 duplicate_rows_df = df[df.duplicated()] #"duplicated???"
3 print(f'Cantidad de filas duplicadas: {duplicate_rows_df.id.count()}')
4
5 #eliminar duplicados
6 df = df.drop_duplicates()
7 #print (df.head())
8
9 # Filas despues de eliminar duplicados
10 print(f'Final: {df.id.count()} filas')
11
```

```
Original: 1018 filas
Cantidad de filas duplicadas: 18
Final: 1000 filas
```

Detect outliers

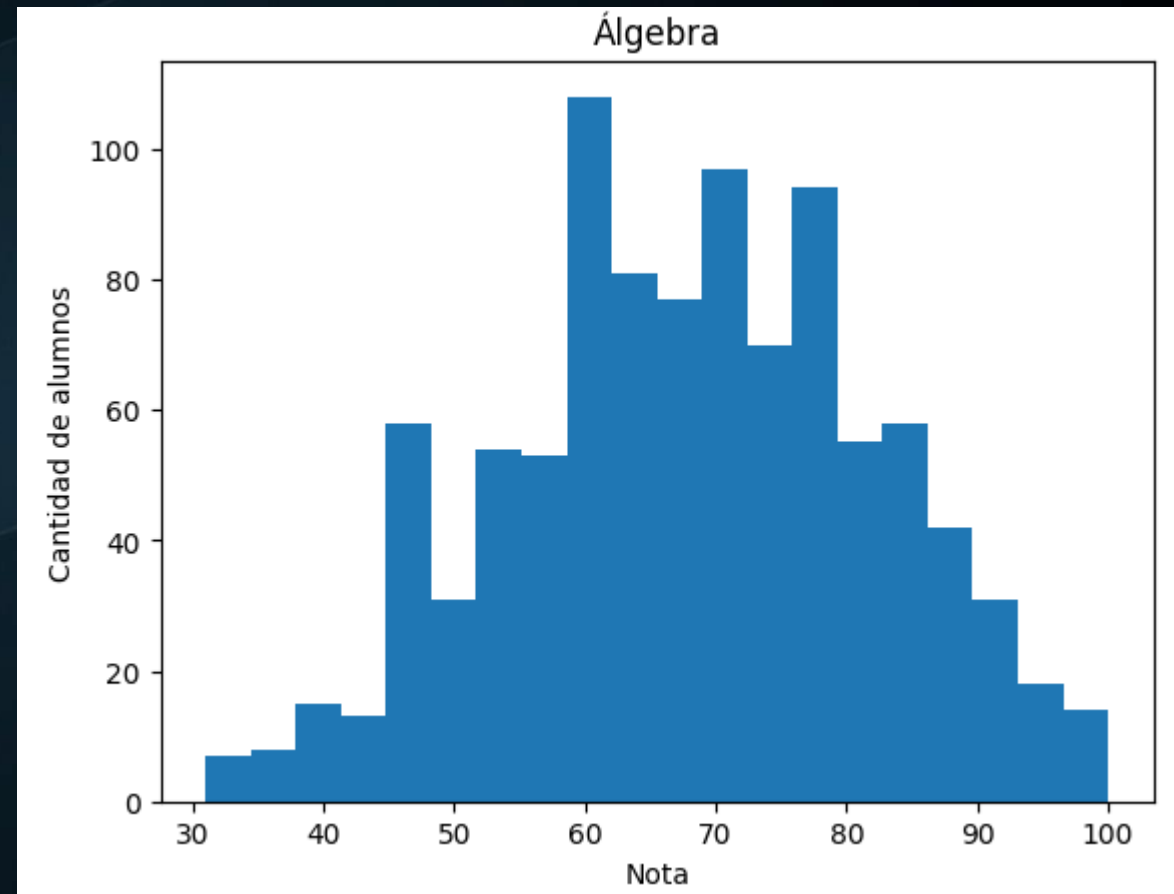
```
#materias renombradas:  
#"math score": "Nota de Matemática",  
#"physics score": "Nota de Física",  
#"chemistry score": "Nota de Química",  
#"algebra_score": "Nota de Álgebra"
```

```
sns.boxplot(x=df['Nota de Matemática'])  
plt.show()  
sns.boxplot(x=df['Nota de Física'])  
plt.show()  
sns.boxplot(x=df['Nota de Química'])  
plt.show()  
sns.boxplot(x=df['Nota de Álgebra'])  
plt.show()
```



Frequencies: Plot histograms

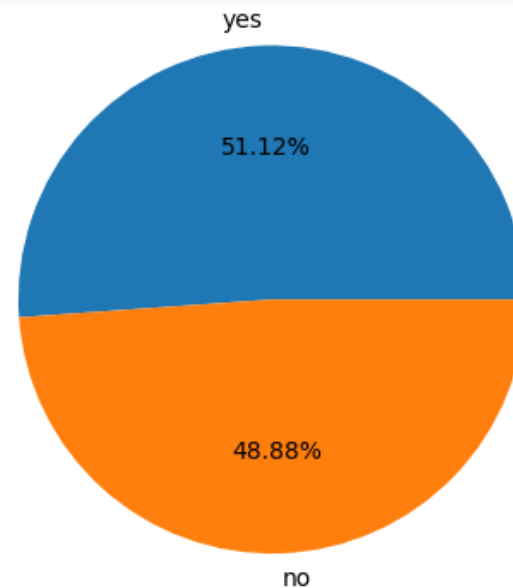
```
1  # "math score": "Nota de Matemática",
2  # "physics score": "Nota de Física",
3  # "chemistry score": "Nota de Química",
4  # "algebra_score": "Nota de Álgebra"
5
6  #Algebra
7  plt.hist(df['Nota de Álgebra'], bins=20)
8  plt.title("Álgebra")
9  plt.ylabel("Cantidad de alumnos")
10 plt.xlabel("Nota")
11 plt.show()
12
13 #Chemistry
14 plt.hist(df['Nota de Química'], bins=20)
15 plt.title("Química")
16 plt.ylabel("Cantidad de alumnos")
17 plt.xlabel("Nota")
18 plt.show()
19
20 #Math
21 plt.hist(df['Nota de Matemática'], bins=20)
22 plt.title("Matemática")
23 plt.ylabel("Cantidad de alumnos")
24 plt.xlabel("Nota")
25 plt.show()
26
27 #Physics
28 plt.hist(df['Nota de Física'], bins=20)
29 plt.title("Física")
30 plt.ylabel("Cantidad de alumnos")
31 plt.xlabel("Nota")
32 plt.show()
33
34
```



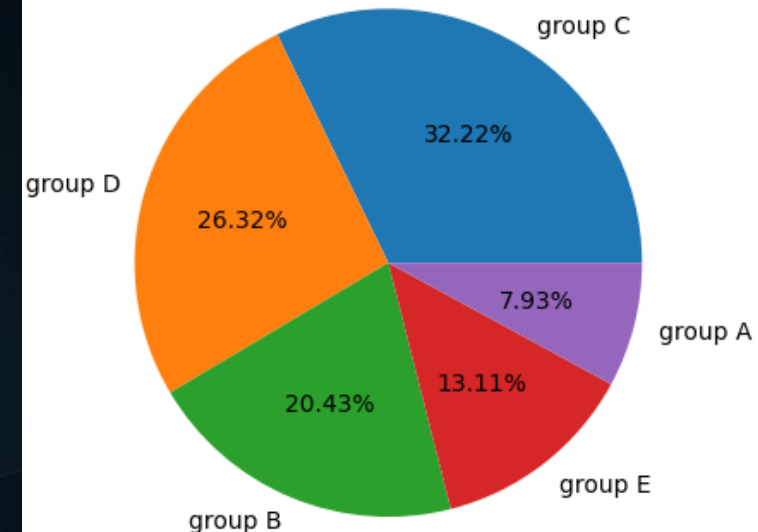
Categorical Values Exploration

```
1 #pandas.value_counts()--> devuelve una serie con valores únicos
2
3 #Nombres actualizados:
4 # "gender": "Género",
5 # "race/ethnicity": "Etnia",
6 # "parental level of education": "Educación de los padres",
7 # "lunch": "Almuerzo",
8 # "employed": "Empleado/a",
9 # "test preparation course": "Curso preparatorio",
10
11 #torta: 'Género'
12 labels = df['Género'].value_counts().index
13 sizes = df['Género'].value_counts()
14 plt.pie(sizes, labels=labels, autopct='%1.2f%%')
15 plt.title('Género')
16 plt.show()
17
18 #torta: 'Etnia'
19 labels = df['Etnia'].value_counts().index
20 sizes = df['Etnia'].value_counts()
21 plt.pie(sizes, labels=labels, autopct='%1.2f%%')
22 plt.title('Etnia')
23 plt.show()
24
25 #torta: 'Empleado/a'
26 labels = df['Empleado/a'].value_counts().index
27 sizes = df['Empleado/a'].value_counts()
28 plt.pie(sizes, labels=labels, autopct='%1.2f%%')
29 plt.title('¿Está empleado?')
30 plt.show()
31
32 #torta: 'Curso preparatorio'
33 labels = df['Curso preparatorio'].value_counts().index
34 sizes = df['Curso preparatorio'].value_counts()
35 plt.pie(sizes, labels=labels, autopct='%1.2f%%')
36 plt.title('¿Tomó el curso preparatorio?')
37 plt.show()
```

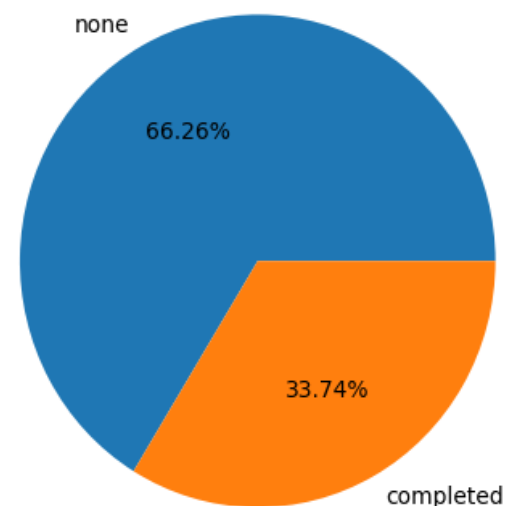
Employed?



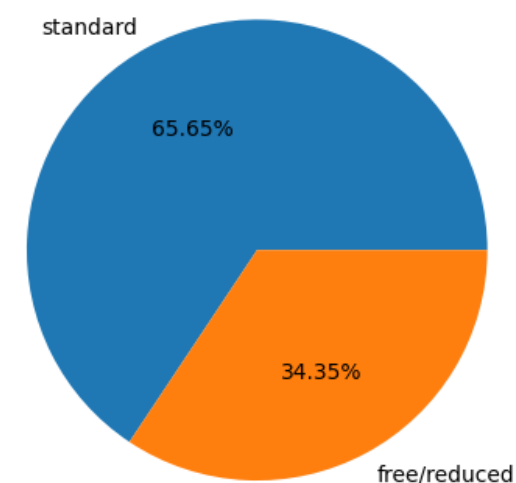
Ethnic Groups



Prep Course Taken



Lunch Type



Christian Monteferrante, MSc., CSCS



cristian.monteferrante@gmail.com



/christian-monteferrante

