

Part 1: Summary and Group Deliverables

Screenshots

```
source /Users/emilygallegos/Downloads/507_groupproject_2025/.venv/bin/activate

The default interactive shell is now zsh.
To update your account to use zsh, please run `chsh -s /bin/zsh`.
For more details, please visit https://support.apple.com/kb/HT208050.
● MacBook-Pro:507_groupproject_2025 emilygallegos$ source /Users/emilygallegos/Downloads/507_groupproject_2025/.venv/bin/activate
● (.venv) MacBook-Pro:507_groupproject_2025 emilygallegos$ python metric_finding.py
```

	playername	team	timestamp	device	metric	value
0	PLAYER_1175	Team: Stony Brook	timestamp	hawkins	Avg. Braking Force(N)	1546.3816
1	PLAYER_1175	Team: Stony Brook	timestamp	hawkins	Avg. Braking Power(W)	-1099.2175
2	PLAYER_1175	Team: Stony Brook	timestamp	hawkins	Avg. Braking Velocity(m/s)	-0.7630
3	PLAYER_1175	Team: Stony Brook	timestamp	hawkins	Avg. Landing Force(N)	1251.3600
4	PLAYER_1175	Team: Stony Brook	timestamp	hawkins	Avg. Propulsive Force(N)	1887.0421
5	PLAYER_1175	Team: Stony Brook	timestamp	hawkins	Avg. Propulsive Power(W)	2667.9769
6	PLAYER_1175	Team: Stony Brook	timestamp	hawkins	Avg. Propulsive Velocity(m/s)	1.4585
7	PLAYER_1175	Team: Stony Brook	timestamp	hawkins	Avg. Relative Propulsive Force(%)	197.6773
8	PLAYER_1175	Team: Stony Brook	timestamp	hawkins	Avg. Relative Propulsive Power(W/kg)	27.4174
9	PLAYER_1175	Team: Stony Brook	timestamp	hawkins	Braking Net Impulse(N.s)	122.0758

```
metric metric count
```

Figure 1. Emily accessing the database using the provided credentials.

```
PS C:\Users\Huma\finalprojecthha507\507_groupproject_2025> & C:/Users/Huma/AppData/Local/Programs/Python/Python313/python.exe

>>> load_dotenv('test.env')
True
>>> load_dotenv('test.env', override=True)
True
>>> sql_username = os.getenv('DB_USER')
>>> sql_password = os.getenv('DB_PASSWORD')
>>> sql_host = os.getenv('DB_HOST')
>>> sql_database = os.getenv('DB_NAME')
>>> sql_username
'ahistudent'
>>> url_string = f"mysql+pymysql://{sql_username}:{sql_password}@{sql_host}:3306/{sql_database}"
>>> conn = create_engine(url_string)
>>> sql_toexecute = """
... select *
... from research_experiment_refactor_test
... limit 50;
... """
>>>
>>> response = pd.read_sql(sql_toexecute, conn)
>>> response.head()
```

	id	playername	timestamp	device	...	session_description	function_description	data_source	created_at
0	1	PLAYER_1175	2018-10-15 19:27:41	hawkins	...	None	None	hawkins	2025-10-21 16:46:54
1	2	PLAYER_1175	2018-10-15 19:27:41	hawkins	...	None	None	hawkins	2025-10-21 16:46:54
2	3	PLAYER_1175	2018-10-15 19:27:41	hawkins	...	None	None	hawkins	2025-10-21 16:46:54
3	4	PLAYER_1175	2018-10-15 19:27:41	hawkins	...	None	None	hawkins	2025-10-21 16:46:54
4	5	PLAYER_1175	2018-10-15 19:27:41	hawkins	...	None	None	hawkins	2025-10-21 16:46:54

Figure 2. Huma accessing the database using the provided credentials.

```
>>> response
```

	id	playername	timestamp	device	metric	value	team	session_type	session_description	function_description	data_source	created_at
0	1	PLAYER_1175	2018-10-15 19:27:41	hawkins	Avg. Braking Force(N)	1546.3816	Team: Stony Brook	None	None	None	hawkins	2025-10-21 16:46:54
1	2	PLAYER_1175	2018-10-15 19:27:41	hawkins	Avg. Braking Power(W)	-1099.2175	Team: Stony Brook	None	None	None	hawkins	2025-10-21 16:46:54
2	3	PLAYER_1175	2018-10-15 19:27:41	hawkins	Avg. Braking Velocity(m/s)	-0.7630	Team: Stony Brook	None	None	None	hawkins	2025-10-21 16:46:54
3	4	PLAYER_1175	2018-10-15 19:27:41	hawkins	Avg. Landing Force(N)	1251.3600	Team: Stony Brook	None	None	None	hawkins	2025-10-21 16:46:54
4	5	PLAYER_1175	2018-10-15 19:27:41	hawkins	Avg. Propulsive Force(N)	1887.0421	Team: Stony Brook	None	None	None	hawkins	2025-10-21 16:46:54
5	6	PLAYER_1175	2018-10-15 19:27:41	hawkins	Avg. Propulsive Power(W)	2667.9769	Team: Stony Brook	None	None	None	hawkins	2025-10-21 16:46:54
6	7	PLAYER_1175	2018-10-15 19:27:41	hawkins	Avg. Propulsive Velocity(m/s)	1.4585	Team: Stony Brook	None	None	None	hawkins	2025-10-21 16:46:54
7	8	PLAYER_1175	2018-10-15 19:27:41	hawkins	Avg. Relative Propulsive Force(%)	197.6773	Team: Stony Brook	None	None	None	hawkins	2025-10-21 16:46:54
8	9	PLAYER_1175	2018-10-15 19:27:41	hawkins	Avg. Relative Propulsive Power(W/kg)	27.4174	Team: Stony Brook	None	None	None	hawkins	2025-10-21 16:46:54
9	10	PLAYER_1175	2018-10-15 19:27:41	hawkins	Braking Net Impulse(N.s)	122.0758	Team: Stony Brook	None	None	None	hawkins	2025-10-21 16:46:54

```
>>> ||
```

Figure 3. Patrick accessing the database using the provided credentials.

```
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS JUPYTER
(.venv) izzitopizzito@Ezzahs-MacBook-Air 507_groupproject_2025 % python3 part1_exploration_clean.py

(.venv) izzitopizzito@Ezzahs-MacBook-Air 507_groupproject_2025 % python3 part1_exploration_clean.py

Loaded ENV values:
DB_HOST: shtn-fallprev.mysql.database.azure.com
DB_USER: ahstudent
DB_NAME: sbu_athletics
DB_TABLE: research_experiment_refactor_test

=== FIRST 10 ROWS (SCREENSHOT THIS) ===

id  playername  timestamp  device  metric  ...  session_type  session_description  function_description  data_source  created_at
0  1  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Braking Force(N)  ...  None  None  None  hawkins  2025-10-21 16:46:54
1  2  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Braking Power(W)  ...  None  None  None  hawkins  2025-10-21 16:46:54
2  3  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Braking Velocity(m/s)  ...  None  None  None  hawkins  2025-10-21 16:46:54
3  4  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Landing Force(N)  ...  None  None  None  hawkins  2025-10-21 16:46:54
4  5  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Propulsive Force(N)  ...  None  None  None  hawkins  2025-10-21 16:46:54
5  6  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Propulsive Power(W)  ...  None  None  None  hawkins  2025-10-21 16:46:54
6  7  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Propulsive Velocity(m/s)  ...  None  None  None  hawkins  2025-10-21 16:46:54
7  8  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Relative Propulsive Force(%)  ...  None  None  None  hawkins  2025-10-21 16:46:54
8  9  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Relative Propulsive Power(W/kg)  ...  None  None  None  hawkins  2025-10-21 16:46:54
9  10  PLAYER_1175  2018-10-15 19:27:41  hawkins  Braking Net Impulse(N.s)  ...  None  None  None  hawkins  2025-10-21 16:46:54

[10 rows x 12 columns]

=== SUCCESS: PART 1 DATABASE ACCESS COMPLETE ===
(.venv) izzitopizzito@Ezzahs-MacBook-Air 507_groupproject_2025 %
```

Figure 4. Ezzah accessing the database using the provided credentials

```
... collecting pymysql
Downloading pymysql-1.1.2-py3-none-any.whl.metadata (4.3 kB)
Downloading pymysql-1.1.2-py3-none-any.whl (45 kB)
Installing collected packages: pymysql
Successfully installed pymysql-1.1.2

id  playername  timestamp  device  metric  value  team  session_type  session_description  function_description  data_source  created_at
0  1  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Braking Force(N)  1546.3816  Team: Stony Brook  None  None  None  hawkins  2025-10-21 16:46:54
1  2  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Braking Power(W)  -1099.2175  Team: Stony Brook  None  None  None  hawkins  2025-10-21 16:46:54
2  3  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Braking Velocity(m/s)  -0.7630  Team: Stony Brook  None  None  None  hawkins  2025-10-21 16:46:54
3  4  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Landing Force(N)  1251.3600  Team: Stony Brook  None  None  None  hawkins  2025-10-21 16:46:54
4  5  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Propulsive Force(N)  1887.0421  Team: Stony Brook  None  None  None  hawkins  2025-10-21 16:46:54
5  6  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Propulsive Power(W)  2667.9769  Team: Stony Brook  None  None  None  hawkins  2025-10-21 16:46:54
6  7  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Propulsive Velocity(m/s)  1.4585  Team: Stony Brook  None  None  None  hawkins  2025-10-21 16:46:54
7  8  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Relative Propulsive Force(%)  197.6773  Team: Stony Brook  None  None  None  hawkins  2025-10-21 16:46:54
8  9  PLAYER_1175  2018-10-15 19:27:41  hawkins  Avg. Relative Propulsive Power(W/kg)  27.4174  Team: Stony Brook  None  None  None  hawkins  2025-10-21 16:46:54
9  10  PLAYER_1175  2018-10-15 19:27:41  hawkins  Braking Net Impulse(N.s)  122.0758  Team: Stony Brook  None  None  None  hawkins  2025-10-21 16:46:54
Columns in table: ['id', 'playername', 'timestamp', 'device', 'metric', 'value', 'team', 'session_type', 'session_description', 'function_description', 'data_source', 'created_at']
```

Figure 5. Avion accessing the database using the provided credentials on a Google Colab notebook.

Summary of Part 1

There are 6617426 unique athletes found in the database.

There are 92 different sports/teams represented in this dataset.

The range of this dataset is from 2018-10-15 19:27:41 to 2025-10-21 12:24:21.

Kinexon is the data source with the most records: kinexon (4073754 records)

There are 0 athletes with missing data in this dataset.

541 athletes have data from multiple sources (2 or 3 systems).

Top 10 common metrics for Hawkins data:

0	System Weight(N)	32373
1	Peak Velocity(m/s)	32123
2	Propulsive Phase(s)	32123
3	Avg. Relative Propulsive Force(%)	32123
4	Avg. Propulsive Power(W)	32123
5	Avg. Propulsive Force(N)	32123
6	Propulsive Net Impulse(N.s)	32123
7	Jump Height(m)	32123
8	Peak Propulsive Force(N)	32123
9	Peak Propulsive Power(W)	32123

Top 10 common metrics for Kinexon data:

0	distance_total	40803
1	accel_load_accum	40803
2	event_count_exertion_category3	40803
3	event_count_exertion_avg_per_minute	40803
4	event_count_exertion_category1	40803
5	event_count_exertion_category2_avg_per_minute	40803
6	event_count_exertion_category3_avg_per_minute	40803
7	event_count_exertion_category2	40803
8	event_count_exertion	40803
9	event_count_exertion_category1_avg_per_minute	40803

Top 10 common metrics for Vald data:

	metric	metric_count
0	leftAvgForce	4275
1	leftCalibration	4275
2	leftImpulse	4275
3	leftMaxForce	4275
4	leftRepetitions	4275
5	leftTorque	4275
6	rightAvgForce	4275
7	rightCalibration	4275
8	rightImpulse	4275
9	rightMaxForce	4275

548 unique metrics exist across all data sources.