

Pràctica 1 (25% nota final)

1. **Context.** Explicar en quin context s'ha recollit la informació. Explicar per què el lloc web triat proporciona aquesta informació.

El context escollit són les estadístiques en l'àmbit del futbol professional a la primera divisió espanyola. La web escollida ha estat <https://playerstats.football>, ja que aquesta proporciona dades en un format únic i fàcil de manipular per part d'una eina d'scraping. Tot i això, el lloc web ens ha presentat amb un seguit de dificultats que hem hagut de gestionar, com ara el fet de que per visualitzar el conjunt de partits disputats per cada equip fos necessari dirigir-se a una nova redirecció.

2. **Títol.** Definir un títol que sigui descriptiu pel dataset.

LaLigaStats.

3. **Descripció del dataset.** Desenvolupar una descripció breu del conjunt de dades que s'ha extret. És necessari que aquesta descripció tingui sentit amb el títol escollit.

Es tracta d'un data set relacionat amb els equips de la primera divisió espanyola. Cada valor ha estat calculat com la mitjana dels últims 30 partits jugats, per les següents estadístiques:

- Possession
- Passes
- Tackles
- Corners
- Shots - Total
- Shots - On target
- Shots - Off target
- Shots - Blocked
- Shots - Outside Box
- Shots - Inside Box
- Fouls
- Offsides
- Yellow Card
- Red Card
- Penalties

Les dades han estat obtingudes de: <https://playerstats.football>

4. **Representació gràfica.** Dibuixar un esquema o diagrama que identifiqui el dataset visualment i el projecte escollit.

La web de partida per la realització del web scraping ha estat la següent:

FC BARCELONA STATS		<div> <div>Passes</div> <div>Tackles</div> <div>Shots on Target</div> <div>Total Shots</div> <div>Yellow Cards</div> <div>Red Cards</div> <div>Fouls Drawn</div> <div>Fouls Committed</div> </div>		<div> <div>Goals Assists</div> <div>HR Woodwork</div> <div>Offsides</div> <div>Dispossessed</div> <div>Accurate Passes</div> <div>Team Stats</div> </div>	
FC Barcelona totals per fixture		Competitions: <input checked="" type="checkbox"/> La Liga <input checked="" type="checkbox"/> Champions League <input checked="" type="checkbox"/> Copa Del Rey <input checked="" type="checkbox"/> Trofeo Joan Gamper		Venue: <input checked="" type="checkbox"/> Home <input checked="" type="checkbox"/> Away <input checked="" type="checkbox"/> Neutral	
<div> <div>FC Barcelona</div> <div> <div>1-1</div> <div>1-0</div> <div>1-2</div> <div>1-0</div> <div>3-1</div> <div>2-0</div> <div>3-0</div> <div>3-0</div> <div>0-0</div> <div>1-1</div> <div>0-3</div> <div>2-1</div> <div>1-1</div> <div>4-2</div> <div>3-0</div> <div>0-1</div> <div>1-2</div> <div>3-3</div> <div>0-0</div> <div>2-3</div> <div>1-2</div> <div>1-2</div> <div>5-2</div> <div>0-4</div> <div>2-1</div> <div>1-0</div> <div>1-6</div> <div>4-1</div> <div>1-1</div> <div>0-2</div> </div> </div>		<div> <div>Possession</div> <div>80%</div> <div>69%</div> <div>50%</div> <div>62%</div> <div>61%</div> <div>71%</div> <div>60%</div> <div>70%</div> <div>68%</div> <div>77%</div> <div>48%</div> <div>67%</div> <div>67%</div> <div>62%</div> <div>61%</div> <div>68%</div> <div>63%</div> <div>61%</div> <div>59%</div> <div>68%</div> <div>81%</div> <div>57%</div> <div>80%</div> <div>77%</div> <div>69%</div> <div>68%</div> <div>69%</div> <div>69%</div> <div>73%</div> <div>64%</div> </div>		<div> <div>Passes</div> <div>811</div> <div>588</div> <div>457</div> <div>708</div> <div>532</div> <div>728</div> <div>524</div> <div>686</div> <div>637</div> <div>638</div> <div>537</div> <div>685</div> <div>589</div> <div>540</div> <div>701</div> <div>726</div> <div>628</div> <div>595</div> <div>777</div> <div>869</div> <div>593</div> <div>938</div> <div>815</div> <div>707</div> <div>697</div> <div>748</div> <div>903</div> <div>854</div> <div>723</div> </div>	
<div> <div>Tackles</div> <div>8</div> <div>13</div> <div>24</div> <div>12</div> <div>14</div> <div>25</div> <div>14</div> <div>19</div> <div>14</div> <div>11</div> <div>18</div> <div>13</div> <div>20</div> <div>16</div> <div>15</div> <div>14</div> <div>19</div> <div>8</div> <div>8</div> <div>20</div> <div>11</div> <div>11</div> <div>9</div> <div>11</div> <div>17</div> <div>11</div> <div>15</div> <div>14</div> </div>		<div> <div>Corners</div> <div>9</div> <div>7</div> <div>7</div> <div>11</div> <div>9</div> <div>6</div> <div>5</div> <div>11</div> <div>5</div> <div>12</div> <div>2</div> <div>3</div> <div>2</div> <div>5</div> <div>3</div> <div>7</div> <div>6</div> <div>5</div> <div>1</div> <div>6</div> <div>1</div> <div>3</div> <div>4</div> <div>6</div> <div>4</div> <div>1</div> <div>7</div> <div>9</div> <div>4</div> </div>		<div> <div>Shots - Total</div> <div>16</div> <div>17</div> <div>11</div> <div>11</div> <div>12</div> <div>9</div> <div>9</div> <div>20</div> <div>5</div> <div>17</div> <div>5</div> <div>9</div> <div>11</div> <div>14</div> <div>14</div> <div>7</div> <div>9</div> <div>17</div> <div>7</div> <div>13</div> <div>15</div> <div>15</div> <div>10</div> <div>16</div> <div>19</div> <div>24</div> <div>14</div> <div>14</div> <div>20</div> <div>11</div> </div>	
<div> <div>Shots - On target</div> <div>5</div> <div>3</div> <div>2</div> <div>3</div> <div>5</div> <div>2</div> <div>0</div> <div>12</div> <div>2</div> <div>6</div> <div>1</div> <div>5</div> <div>5</div> <div>9</div> <div>8</div> <div>3</div> <div>2</div> <div>5</div> <div>2</div> <div>4</div> <div>3</div> <div>5</div> <div>8</div> <div>9</div> <div>4</div> <div>10</div> <div>10</div> <div>6</div> <div>10</div> <div>4</div> </div>		<div> <div>Shots - Off target</div> <div>11</div> <div>14</div> <div>9</div> <div>8</div> <div>7</div> <div>7</div> <div>9</div> <div>8</div> <div>3</div> <div>11</div> <div>4</div> <div>4</div> <div>6</div> <div>5</div> <div>6</div> <div>4</div> <div>7</div> <div>12</div> <div>5</div> <div>9</div> <div>12</div> <div>10</div> <div>2</div> <div>7</div> <div>15</div> <div>14</div> <div>4</div> <div>8</div> <div>10</div> <div>7</div> </div>		<div> <div>Shots - Blocked</div> <div>3</div> <div>4</div> <div>3</div> <div>1</div> <div>3</div> <div>1</div> <div>2</div> <div>3</div> <div>0</div> <div>2</div> <div>2</div> <div>1</div> <div>0</div> <div>2</div> <div>1</div> <div>4</div> <div>3</div> <div>0</div> <div>0</div> <div>0</div> <div>4</div> <div>0</div> <div>0</div> <div>6</div> <div>7</div> <div>4</div> <div>2</div> <div>6</div> <div>6</div> </div>	
<div> <div>Shots - Outside Box</div> <div>10</div> <div>4</div> <div>3</div> <div>2</div> <div>5</div> <div>6</div> <div>2</div> <div>8</div> <div>2</div> <div>4</div> <div>3</div> <div>2</div> <div>1</div> <div>2</div> <div>4</div> <div>8</div> <div>5</div> <div>4</div> <div>2</div> <div>3</div> <div>4</div> <div>7</div> <div>3</div> <div>2</div> <div>3</div> <div>7</div> <div>15</div> <div>6</div> <div>7</div> <div>7</div> </div>		<div> <div>Shots - Inside Box</div> <div>6</div> <div>12</div> <div>9</div> <div>9</div> <div>7</div> <div>3</div> <div>6</div> <div>12</div> <div>4</div> <div>13</div> <div>2</div> <div>5</div> <div>8</div> <div>11</div> <div>3</div> <div>13</div> <div>12</div> <div>8</div> <div>11</div> <div>9</div> <div>12</div> <div>7</div> <div>11</div> <div>11</div> <div>10</div> <div>8</div> <div>5</div> <div>14</div> <div>6</div> </div>		<div> <div>Fouls</div> <div>8</div> <div>11</div> <div>16</div> <div>17</div> <div>15</div> <div>9</div> <div>14</div> <div>11</div> <div>9</div> <div>14</div> <div>9</div> <div>16</div> <div>9</div> <div>11</div> <div>11</div> <div>15</div> <div>9</div> <div>5</div> <div>8</div> <div>14</div> <div>6</div> <div>11</div> <div>7</div> <div>13</div> <div>12</div> <div>12</div> <div>7</div> <div>8</div> <div>13</div> </div>	
<div> <div>Offsides</div> <div>1</div> <div>1</div> <div>2</div> <div>2</div> <div>3</div> <div>4</div> <div>0</div> <div>1</div> <div>1</div> <div>2</div> <div>3</div> <div>2</div> <div>1</div> <div>3</div> <div>3</div> <div>0</div> <div>1</div> <div>3</div> <div>1</div> <div>2</div> <div>6</div> <div>1</div> <div>2</div> <div>1</div> <div>1</div> <div>2</div> <div>2</div> <div>1</div> <div>3</div> </div>		<div> <div>Yellow Card</div> <div>0</div> <div>2</div> <div>1</div> <div>1</div> <div>0</div> <div>2</div> <div>1</div> <div>5</div> <div>4</div> <div>3</div> <div>1</div> <div>3</div> <div>2</div> <div>2</div> <div>1</div> <div>3</div> <div>2</div> <div>3</div> <div>1</div> <div>2</div> <div>0</div> <div>3</div> <div>2</div> <div>0</div> <div>4</div> <div>3</div> <div>2</div> <div>0</div> <div>3</div> </div>		<div> <div>Red Card</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>1</div> <div>0</div> <div>0</div> <div>1</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>1</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> </div>	
<div> <div>Penalties</div> <div>0</div> <div>1</div> <div>0</div> <div>0</div> <div>1</div> <div>0</div> <div>0</div> <div>1</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>1</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>1</div> <div>0</div> </div>		<div> <div>Throw-ins</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> </div>		<div> <div>Free Kicks</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> </div>	
<div> <div>Goal Kicks</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> </div>		<div> <div>Goal Kicks</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> </div>		<div> <div>Goal Kicks</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> <div>0</div> </div>	

El format resultant buscat, respon al següent esquema:

Name	Team
Fouls	##
Corners	##
Offsides	##
Passes	##
Penalties	##
Possession	##
Red Card	##
Shots - Blocked	##
Shots - Inside Box	##
Shots - Off target	##
Shots - On target	##
Shots - Outside Box	##
Shots - Total	##
Tackles	##
Yellow Card	##
Passes	##

De la manera que el resultat final, ha estat el següent:

Name	Fouls	Corners	Offsides	Passes	Penalties	Possession	Red Card	Shots - Blocked	Shots - Inside Box	Shots - Off target	Shots - On target	Shots - Outside Box	Shots - Total	Tackles	Yellow Card	Passes
real-madrid	1.0E+16	6.0E+15	2.3E+16	5.8E+15	1.0E+00	5.8E+02	3.3E+15	2.9E+01	9.8E+01	1.0E+16	5.4E+01	6.0E+01	1.6E+16	1.5E+16	1.7E+16	
real-sociedad	1.4E+02	4.4E+02	1.8E+01	4.6E+04	2.8E+01	5.5E+03	4.0E+00	2.0E+02	8.0E+01	7.1E+02	4.3E+02	3.4E+02	1.1E+02	1.5E+02	2.2E+02	
sevilla	1.2E+16	4.5E+15	1.7E+16	5.5E+15	2.0E+00	6.1E+16	1.3E+16	1.5E+16	8.0E+15	8.2E+01	4.4E+01	4.8E+15	1.3E+02	1.2E+16	2.4E+01	
athletic-madrid	1.2E+16	5.8E+15	2.8E+01	4.8E+03	1.0E+00	5.2E+02	1.3E+16	1.9E+16	8.7E+15	8.4E+01	4.8E+15	4.6E+15	1.3E+16	1.6E+16	2.9E+16	
rayo-vallecano	1.4E+16	3.6E+16	2.7E+15	3.5E+03	1.8E+16	4.7E+15	9.1E+15	2.0E+01	6.8E+15	8.0E+01	3.8E+16	5.0E+01	1.2E+16	1.5E+16	2.6E+16	
real-betis	1.3E+16	4.9E+01	1.9E+01	4.7E+15	3.0E+00	5.4E+02	1.3E+16	2.3E+16	7.8E+01	8.4E+15	4.7E+01	5.3E+01	1.3E+16	1.6E+16	2.7E+16	
osasuna	1.2E+16	3.9E+16	1.8E+16	3.6E+15	2.7E+16	4.5E+16	3.3E+15	1.4E+16	6.5E+15	7.0E+01	3.3E+16	3.9E+16	1.0E+16	1.4E+16	2.1E+15	
athletic-club	1.2E+16	6.1E+15	1.9E+15	3.8E+15	1.6E+16	4.6E+02	1.1E+16	1.7E+16	8.0E+01	7.8E+15	3.7E+15	3.6E+16	1.2E+16	1.2E+16	1.6E+15	
fc-barcelona	1.2E+16	6.7E+15	1.9E+16	6.2E+15	2.1E+16	6.5E+15	2.1E+16	1.9E+16	7.6E+15	7.6E+15	4.3E+15	3.9E+15	1.2E+16	1.6E+02	2.0E+01	
valencia	9.7E+15	4.4E+15	1.8E+01	3.8E+16	2.3E+16	4.7E+15	1.3E+16	1.5E+01	5.9E+15	7.0E+01	3.4E+01	4.8E+15	1.1E+16	1.2E+16	2.4E+15	
espanyol	1.1E+16	6.2E+15	2.1E+16	3.7E+04	1.7E+16	4.6E+16	1.7E+16	1.6E+16	6.4E+15	7.8E+15	4.2E+15	5.3E+15	1.2E+02	1.4E+03	2.1E+16	
mallorca	1.3E+02	4.0E+01	2.0E+01	3.3E+15	0.0E+00	4.5E+16	2.5E+01	1.1E+16	5.6E+15	6.7E+15	4.3E+15	5.3E+02	1.1E+02	1.6E+16	2.8E+16	
villarreal	1.2E+02	4.5E+01	2.4E+15	5.1E+16	0.0E+00	5.9E+02	1.4E+16	2.1E+15	8.9E+15	7.4E+15	4.3E+15	2.5E+01	1.2E+16	1.5E+16	2.8E+16	
celta-de-vigo	1.5E+02	3.5E+15	1.9E+16		2.0E+00	5.6E+16	1.3E+16	1.1E+01	6.3E+01	6.1E+01	3.5E+01	3.3E+16	9.6E+01	1.8E+02	2.8E+01	4.8E+16
elche	1.2E+16	4.4E+01	1.6E+16	3.9E+16	3.3E+15	4.7E+02	6.7E+15	1.2E+01	5.2E+15	5.5E+01	2.9E+16	3.0E+15	8.4E+15	1.4E+02	2.6E+16	
deportivo-alaves	1.3E+16	4.3E+15	1.7E+16	3.4E+16	1.3E+16	4.3E+15	1.0E+00	1.1E+01	6.6E+15	6.2E+15	3.0E+15	3.1E+16	9.3E+15	1.3E+02	2.3E+16	
cadix	9.7E+15	4.0E+15	1.5E+01	2.7E+15	2.7E+16	3.4E+16	6.7E+15	6.7E+15	5.4E+15	5.7E+01	2.3E+01	2.8E+16	8.0E+01	1.6E+16	2.5E+15	
granada	1.3E+16	4.2E+15	1.9E+16	3.2E+16	1.3E+16	4.2E+15	6.7E+15	1.4E+16	5.1E+15	5.9E+01	2.9E+15	4.1E+15	9.0E+15	1.4E+16	3.2E+16	
levantante	1.4E+16	3.8E+16	1.7E+16	4.0E+05	2.1E+16	4.8E+16	4.2E+16	1.3E+02	5.3E+15	6.1E+15	2.8E+16	3.4E+16	8.9E+03	1.4E+16	2.3E+16	
getafe	1.5E+02	3.9E+15	1.7E+16	3.1E+15	1.3E+16	4.2E+15	1.7E+16	1.8E+16	5.8E+15	7.4E+01	2.6E+01	4.5E+15	1.0E+02	1.6E+16	3.0E+15	

5. **Contingut.** Explicar els camps que inclou el dataset, el període de temps de les dades i com s'ha recollit.

1. Name. Es tracta del nom de l'equip.
2. Possession. Es tracta de la possessió mitja en els últims 30 partits.
3. Passes. Es tracta del nombre mig de passades en els partits analitzats.
4. Tackles. Es tracta del nombre de disputes per la pilota, de mitja durant els últims 30 jocs.
5. Corners. És la mitjana de córners per partit, analitzant-ne els 30 últims.
6. Shots – Total. Es tracta del nombre mig de xuts als últims 30 partits.
7. Shots - On target. Es tracta del nombre mig de xuts a porteria als últims 30 partits.
8. Shots - Off target. Es tracta del nombre mig de xuts a fora als últims 30 partits.
9. Shots – Blocked. Es tracta del nombre mig de xuts aturats als últims 30 partits.
10. Shots - Outside Box. Es tracta del nombre mig de xuts des de fora l'àrea als últims 30 partits.
11. Shots - Inside Box. Es tracta del nombre mig de xuts des de dins de l'àrea als últims 30 partits.
12. Fouls. Es tracta del nombre mig de faltes comeses per l'equip als últims 30 partits.
13. Offsides. Es tracta del nombre mig de fores de joc als últims 30 partits.
14. Yellow Card. Es tracta del nombre mig de targetes grogues als últims 30 partits.
15. Red Card. Es tracta del nombre mig de targetes vermelles als últims 30 partits.
16. Penalties. Es tracta del nombre mig de penaltis als últims 30 partits.

6. **Agraïments.** Presentar el propietari del conjunt de dades. És necessari incloure cites d'anàlisis anteriors o, en cas de no haver-n'hi, justificar aquesta cerca amb anàlisis similars. Justificar quins passos s'han seguit per actuar d'acord amb els principis ètics i legals en el context del projecte.

En primer lloc, a la pàgina web des d'on hem recopilat les estadístiques, PlayerStats (<https://playerstats.football/>). Hem escollit aquesta web per tenir un format en forma tabular, on trobàvem dades per a tots els equips de la primera divisió en diverses competicions.

Des del punt de vista ètic i legal, no hem afrontat grans dilemes ja que l'objectiu de l'exercici era acadèmic, i la intenció de font de dades és precisament oferir una forma fàcil d'estudiar estadístiques futbolístiques.

7. **Inspiració.** Explicar per què és interessant aquest conjunt de dades i quines preguntes es pretenen respondre. És necessari comparar amb les anàlisis anteriors presentades a l'apartat 6.

Aquest conjunt de dades ens resulta interessant perquè és el càlcul de les mitjanes de les diferents estadístiques estudiades, fixant-nos en els últims 30 partits, per aquelles competicions més importants. De manera que ens permet copsar l'estat de forma dels diferents equips de la primera divisió d'una manera objectiva.

8. **Llicència.** Seleccionar una d'aquestes llicències pel dataset resultant i justificar el motiu de la seva selecció:

- Released Under CC0: Public Domain License
- Released Under CC BY-NC-SA 4.0 License
- Released Under CC BY-SA 4.0 License
- Database released under Open Database License, individual contents under Database Contents License
- Other (specified above)
- Unknown License

Hem escollit la llicència Creative Commons 4.0 International, que es correspon a la Released Under CC BY-SA 4.0 License.

El motiu és que es permet la lliure distribució del data set, així com la transformació i modificació per qualsevol propòsit. Al tractar-se d'un data set generat en un context educatiu, ens ha semblat un tipus de llicència adequada, ja que l'objectiu és que qualsevol persona pugui fer estudis sobre aquest.

9. **Codi.** Adjuntar al repositori Git el codi amb el qual s'ha generat el dataset, preferiblement en Python o, alternativament, en R.

https://github.com/mtuneuf/practica1_tipologia/tree/main/Codi

10. **Dataset.** Publicar el dataset obtingut(*) en format CSV a Zenodo amb una breu descripció. Obtenir i adjuntar l'enllaç del DOI.

https://zenodo.org/record/5636156#.YX_Tcp6ZOUk

Contribucions

Contribucions	Signatura
Investigació prèvia	MTF, MSB
Redacció de les respostes	MTF, MSB
Desenvolupament del codi	MTF, MSB