

Arribades a l'Aeroport de Barcelona el Prat

Descripció del dataset

Aquest conjunt de dades conté la informació relativa als vols d'arribada a l'Aeroport del Prat de Barcelona. Es poden veure per un dia determinat (11/11/2019) tots els vols que tenen previst aterrar a l'aeroport en una franja de temps determinada.

Depenent del moment de l'extracció també es pot conèixer l'hora en que ha aterrat i, per tant, calcular les demores o cancel·lacions de tots els vols en funció del seu origen, durada, companyia, etc.

Imatge identificativa

Arrivals			
Time	From	Arrival	Expected
04:05	Los Angeles	AA680	Landolt 15:37
13:10	Geneva	BA229	Landolt 16:42
15:30	Glasgow	AA649	Landolt 15:17
15:31	Birmingham	AA638	Landolt 16:18
15:48	London	AA641	Landolt 16:41
15:56	Edin	BA682	Expected 16:19
16:05	Geneva	BA731	Landolt 16:38
16:08	Dublin	BA681	Landolt 16:19
16:15	Montreal	AH675	Landolt 16:15
16:25	Dublin	BA681	Expected 16:06
16:35	Geneva	BA761	Landolt 17:06
16:48	Manchester	AA626	Landolt 17:23
16:49	Dublin	BA695	Stop-returned
16:52	Dublin	AH630	Stop-returned
16:55	Millan-Shute	BA663	Landolt 17:13
17:00	Athens	BA647	Landolt 17:41
17:10	Australia	BA408	Landolt 17:19
17:40	Copenhagen	BA672	Landolt 17:43
17:55	Stockholm	BA687	Expected 18:32
17:55	Birmingham	AA681	Expected 18:50
17:25	Paris CDG	BA719	Cancelled
17:40	Geneva	BA710	Expected 18:12
17:45	St Petersburg	BA670	Landolt 17:46
18:00	Wolverhampton	AA639	Cancelled
18:00	Vienna	BA689	Expected 18:48
18:10	Edinburgh	AA680	Expected 18:15
18:15	Minster	BA673	Expected 18:36
18:15	East	BA705	Expected 18:19
18:20	Millan-Shute	BA687	Cancelled
18:25	Edin	BA763	Expected 17:59
18:45	London	BA671	Expected 18:25
18:55	Dublin	BA699	Cancelled
19:00	Abingdon	W7561	Cancelled
19:00	Abingdon	W7754	Cancelled
19:05	Luton	BA643	Expected 17:38
19:10	Geneva	BA712	Cancelled
19:15	Prague	BA681	Cancelled
19:25	Düsseldorf	CE748	Cancelled
19:25	Amsterdam	BA645	Cancelled
20:00	Brussels	BA687	Cancelled
20:15	Edin	BA681	Cancelled
20:20	Waltham	BA661	Cancelled
20:25	Vienna	BA705	Expected 18:53
20:35	Dublin	BA687	Expected 19:17
20:40	Edinburgh	BA740	Cancelled
20:45	Paris CDG	BA707	Cancelled
21:00	Glasgow	BA707	Cancelled
21:00	Montreal	BA707	Cancelled
21:05	Amsterdam	BA673	Delayed
21:25	Geneva	BA751	Cancelled
21:30	Athens	BA653	Expected 21:39
21:05	Paris-CDG	BA709	Cancelled
21:10	Oslo	BA769	Cancelled
21:15	Montreal	BA673	Expected 21:48
21:15	Prague	BA687	Cancelled
Monday, March 29th 1999			
00:00	Hong Kong	BA658	Expected 00:24
00:05	Hong Kong	BA653	Expected 00:38
00:10	Johannesburg	BA658	Cancelled
00:20	Hong Kong	BA658	Expected 00:36

Imatge 1. Fotografia dels panells d'informació sobre les arribades, disponibles a les terminals de l'Aeroport de Barcelona.

Context

Aquest projecte fa servir mitjançant un robot de *Web Scrapping* la informació sobre les arribades a l'aeroport de Barcelona. Aquesta [web](#), tal com s'especifica, no és la web oficial de l'aeroport. Fa servir la informació que compra a un proveïdor i la publica per obtenir visites i obtenir ingressos a través d'anuncis.

Aquest lloc web, dona un detall d'informació major que el que donen les webs oficials. Dona detall sobre el model d'avió que arriba i si hi ha alguna demora en les arribades.

Contingut

En aquest conjunt de dades podem trobar:

link: enllaç d'on s'estan obtenint les dades.

code: codi del vol. Identificació internacional que inclou dues lletres corresponents a la companyia i un nombre corresponent a la ruta.

company: nom de la companyia que opera el vol.

status: situació en la que es troba el vol. Indica si ja ha aterrat, si ha estat endarrerit, etc.

plane: model d'avió amb el que es fa el trajecte.

origin: lloc de sortida del vol. En aquest conjunt de dades aquest valor serà el mateix per a tots els registres.

depDate: data de sortida del vol.

depTime: Hora en que surt el vol.

sDepTime: Hora estimada o prevista de sortida.

depTerm: terminal origen.

delGate: Porta origen

destin: aeroport de destinació.

arrDate: data d'arribada del vol.

arrTime: Hora d'arribada del vol.

sArrTime: hora estimada o prevista d'arribada.

arrTerm: terminal de l'aeroport d'arribada.

arrGate: porta de l'aeroport d'arribada.

Les dades recollides són una fotografia del que mostrava la pàgina web a 11/11/2019. Mostra els vols del dia.

Agraïments

Les dades es capturen de la pàgina web *Barcelona-Airpot*. Aquest portal, tot i no ser l'oficial, ofereix una informació més completa i de forma més amigable que la web oficial d'AENA.

Inspiració

Aquesta projecte desperta interès pel que fa a l'ús de la tècnica de Web Scrapping perquè exigeix entrar en diferents directoris i capturar informació a diferents nivells. A més, aquest projecte permet ser ampliat en posteriors fases que incloguin la normalització dels camps i l'automatització del bot perquè entri periòdicament i reculli nova informació així com actualitzi la dels vols que ja han aterrat i si ho han fet segons l'horari previst.

Aquesta idea sorgeix d'una conversa casual en la que es va esmentar aquesta web. No té cap objectiu concret més que l'ús de la tècnica. Tanmateix, el projecte podria respondre preguntes que es planteessin comerços a agents relacionats amb el turisme. Ja que poden disposar de dades a temps reals dels potencials clients que arriben a Barcelona.

El projecte és fàcilment escalable ja que existeix una ampla col·lecció de webs no oficials sobre aeroports d'arreu del món que segueixen el mateix format.

Llicència

La llicència escollida per aquest conjunt de dades i el bot que l'obté és **CC BY-NC-SA 4.0 License**. Donat que l'objectiu d'aquest projecte és estrictament didàctic, s'escull aquesta llicència, per obrir el contingut que es pugui generar.

Adicionalment es demana que es citi la font i permet que es faci ús de la mateixa o de derivades. S'explicita que no es faci un ús comercial, ja que, encara que aquest contingut estigui disponible gratuïtament al web, no en som els propietaris ni ens sembla ètic vendre aquestes dades.

Recursos

- Subirats, L., Calvo, M. (2018). Web Scraping. Editorial UOC.
- Masip, D. El lenguaje Python. Editorial UOC.
- Lawson, R. (2015). Web Scraping with Python. Packt Publishing Ltd. Chapter 2. Scraping the Data.
- Simon Munzert, Christian Rubba, Peter Meißner, Dominic Nyhuis. (2015).
- Automated Data Collection with R: A Practical Guide to Web Scraping and Text Mining. John Wiley & Sons.
- Tutorial de Github <https://guides.github.com/activities/hello-world>