

Tipologia i cicle de vida de les dades - Pràctica 1

Autors:

- Jordi Bellviure Figueras
- Xenia Casanovas Díez

1. Context

La pàgina web EscapeRadar.com conté un registre de totes les sales d'escape room d'Espanya i on els usuaris poden afegir valoracions en funció de les seves experiències. A part, també s'hi pot trobar informació pràctica com dades de contacte, disponibilitat o característiques de la sala (terror, apte per embarassades, per a tota la família, disponible en anglès, etc.).

A partir de la informació extreta d'aquesta pàgina web s'ha construït un conjunt de dades que recull aquesta informació amb la voluntat de fer possible un anàlisi d'aquestes.

2. Títol

Sales d'Escape Room de l'estat espanyol

3. Descripció del dataset

El dataset s'ha extret a partir del nom de la sala i inclou la informació bàsica d'aquesta per tal de facilitar una visió globalitzada de la situació d'aquest tipus de sala en el marc de l'estat espanyol. A més, també inclou característiques de cada una de les sales com per exemple el públic al que va dirigit (edat mínima, si és familiar, de terror, etc.), així com informació de l'adaptabilitat de la sala (si és apta per embarassades, per persones amb diversitat funcional o claustrofòbia, etc.).

En aquelles que ho tenen disponible, s'inclou informació de la disponibilitat de cada una de les sales en el moment de l'extracció de les dades i pels propers 6 dies per tal de poder fer un anàlisi comparativa (en funció de la temàtica, públic, etc.) de la demanda de les sales.

Finalment, s'inclou informació de contacte de l'empresa encarregada de la sala d'escape room, com l'adreça, el telèfon o el correu electrònic, així com l'estat de la sala (oberta, tancada o tancada temporalment).

4. Representació gràfica

Dibuixar un esquema o diagrama que identifiqui el dataset visualment i el projecte escollit.

5. Contingut

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

- id: Identificador de la sala (assignat per nosaltres)
- name: Nom de la sala
- imgName: Nom de la imatge extreta de la sala guardada a la carpeta "escapeImg"
- punctuation: Puntuació general de la sala
- companyName: Nom de la empresa que es propietaria de la sala
- locZone: Nom de la localització de la sala (normalment es correspon amb la ciutat)
- minPlayers: Número de jugadors mínim per jugar la sala
- maxPlayers: Número de jugadors màxims que poden jugar
- timelapse: Temps estimat o temps màxim de realització de la sala
- lowPrice: Rang de preu inferior de la sala
- highPrice: Rang de preu superior de la sala
- difLevel: Nivell de dificultat
- audience
- category
- horror: Puntuació de terror/horror de la sala
- address: Adreça de la sala
- action: Temàtica d'acció. S = sí, N = no.
- adventure: Temàtica d'aventures. S = sí, N = no.
- cifi: Temàtica de ciència ficció. S = sí, N = no.
- childish: Sala infantil. S = apte per a infants, N = no apte per a infants.
- investigation: Temàtica d'investigació. S = sí, N = no.
- scary: Temàtica de terror. S = sí, N = no.
- mystery: Temàtica de misteri. S = sí, N = no.
- eOnline: És una sala que es juga Online.
- eExterior: És una sala que es juga a l'exterior.
- jPortatil: És un joc portàtil.
- VR: És una sala de realitat virtual.
- eHall: És una sala del tipus Escape Hall.
- salavs: La sala ofereix mode VS amb 2 sales iguals.
- empresas: Indicada per a empreses.
- grups: Indicada per a grups.
- family: Indicada per a famílies.
- kids: Indicada per a infants.
- pregnant: Apte per a persones embarassades.
- english: Disponible en anglès.
- funcDiversity: Apte per persones amb diversitat funcional.
- claustrofobia: Apte per persones amb claustrofòbia.
- ptsComenGen: Puntuació atorgada pels usuaris en aspectes de "General".
- ptsComenAmbi: Puntuació atorgada pels usuaris en aspectes de "Ambientació".

- ptsComenEnig: Puntuació atorgada pels usuaris en aspectes de “Enigma”.
- ptsComenInm: Puntuació atorgada pels usuaris en aspectes de “Inmersió”.
- ptsComenHorror: Puntuació atorgada pels usuaris en aspectes de “Horror”.
- availableTimes: Hores disponibles pels propers 6 dies en el moment de l'extracció de les dades.
- unavailableTimes: Hores ocupades pels propers 6 dies en el moment de l'extracció de les dades.
- companyAddress: Adreça postal de l'empresa propietària de la sala.
- companyPhone: Telèfon de contacte de l'empresa propietària de la sala.
- companyEmail: Adreça de correu electrònic de l'empresa propietària de la sala.
- companyWeb: Direcció de la pàgina web de l'empresa propietària de la sala.
- state: Estat de la sala: ABIERTA, CERRADA TEMPORALMENTE, CERRADA.

6. Agraïments

La pàgina web utilitzada “escaperadar.com” es propietat de la empresa “ESCAPE RADAR, S.L.”.

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.

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.

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

9. Codi

Enllaç a GitHub:

[jbellviurefi/-Web-scraping-Escape-Radar \(github.com\)](https://github.com/jbellviurefi/-Web-scraping-Escape-Radar)

10. Dataset

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

11. Vídeo

S'ha de lliurar un vídeo explicatiu de la pràctica on cadascun dels integrants del grup expliqui amb les seves pròpies paraules tant les respostes del projecte com el codi utilitzat per a dur a terme l'extracció. El vídeo ha de ser enviat a través d'un enllaç a Google Drive que heu de proporcionar, juntament amb l'enllaç al repositori Git, al moment de lliurar la pràctica.

ALTRES

- Arxiu robots.txt: <https://www.escaperadar.com/robots.txt>

User-agent: *

Disallow: /admin.php

Disallow: /*?

Disallow: /rooms/

Sitemap: <https://www.escaperadar.com/sitemap.xml>

User-Agent: Googlebot

Allow: /*.css\$

Allow: /*.js\$

- Sitemap: <https://www.escaperadar.com/sitemap.xml>

Estructura general: