



PROJETO FINAL

INSTITUTO POLITÉCNICO DE COIMBRA - ESCOLA SUPERIOR DE EDUCAÇÃO

CURSO LICENCIATURA DE COMUNICAÇÃO E DESIGN MULTIMÉDIA
Programação e Algoritmos II - Docente: João Orvalho

INSTITUTO POLITÉCNICO DE COIMBRA – ESCOLA SUPERIOR DE EDUCAÇÃO
CURSO LICENCIATURA DE COMUNICAÇÃO E DESIGN MULTIMÉDIA
Programação e Algoritmos II - Docente: João Orvalho

Parte I – Contexto

No âmbito da unidade curricular Programação e Algoritmos II do 1.º ano do curso de Licenciatura de Comunicação e Design Multimédia, 2º semestre, sob orientação do docente João Orvalho, foi-nos solicitado a elaboração de um trabalho.

Através da elaboração deste trabalho, pretendemos abordar o tema do Turismo (particularmente em Portugal), uma área que devido á pandemia tem sofrido diversas alterações.

Para o desenvolvimento deste trabalho foi realizada uma pesquisa para a criação de uma base de dados, tendo como suporte os websites referidos na bibliografia.

Para o desenvolvimento deste projeto final visualizamos exemplos práticos mostrados pelo Docente, no entanto com a pesquisa para a composição deste projeto conseguimos explorar e adquirir novos conhecimentos na área da programação.

Parte II – Relatório do Projeto

Capítulo I - Fundamentação dos dados

A recolha de dados devidamente estruturados é crucial para o desenvolvimento do nosso projeto final, no entanto grande parte dos dados encontrados tratam-se de dados com fraca qualidade e construídos indevidamente.

Posto isto, vimo-nos a ser obrigadas a proceder a uma limpeza dos dados, sendo que estes estavam em mau estado. Por exemplo, palavras em campos que deveriam ser numéricos e organização por tabela dos dados.

Sobre o tema escolhido para este projeto foi realmente difícil encontrar dados que estivessem em formato standard, o que seria muito benéfico para o bom desenvolvimento deste trabalho.

Foi utilizando dois formatos de ficheiro, Json e CSV. O Json são representados como pares de valor-chave num formato semiestruturado, este contém estruturas hierárquicas. O formato CSV são considerados semiestruturados, mas não podem, naturalmente, representar dados hierárquicos ou relacionais.

Recolhemos dados através de download mas também experimentamos exportar dados através de link URL, no entanto este não disponíveis no projeto.

No caso dos dados “Ano de Atendimento” estes foram importados através de sistema URL, o formato dos dados fazia com que fosse impossível a utilização dos mesmo. Para solucionar o nosso problema passamos á criação de um CSV standard com os dados que queríamos utilizar corretamente organizados.

Deparamo-nos com a existência de valores “ausentes”, sendo que estes eram bastante importantes para os objetivos do projeto, conseguimos resolver com auxílio de outros exercícios.

Para a elaboração da nossa base de dados, importamos ficheiros em CSV e Json, no entanto muitos dos dados encontrados estavam em outros formatos, o que criou alguns problemas. Como por exemplo xlsx, para conseguirmos utilizar estes dados com base no ficheiro "copia de despesas media de turistas" criamos uma tabela ("preços") com os dados e não importamos diretamente o ficheiro.

INSTITUTO POLITÉCNICO DE COIMBRA – ESCOLA SUPERIOR DE EDUCAÇÃO**CURSO LICENCIATURA DE COMUNICAÇÃO E DESIGN MULTIMÉDIA**

Programação e Algoritmos II - Docente: João Orvalho

Deparamo-nos com a falta de dados para a execução de exercícios relativos á longitude e á latitude, por isso importamos mais dados que não pertencem só a Portugal.

Em relação ás considerações éticas foi seguido sempre com rigor as indicações presentes nos dados.

Capítulo II - Aplicação do Repositório**Dormidas nº estadia.json**

- Aceder ao tamanho da tabela de dados de 2019
- Renomear colunas e mostrar colunas importantes
- Quantidade de posições com valores em falta
- Eliminar colunas com dados em falta
- Aceder aos dados que pertencem a Portugal no ano de 2019
- Eliminar linhas com dados em falta

alojamentos2019.csv

- Total de informação em falta e posições em que se encontram
- Ordenação por ordem alfabética dos nomes dos alojamentos
- Identificar as moradias que se enquadram na categoria de motel
- Localização do motel Alameda

Ano atendimento.csv

- Visualização do ficheiro como objeto
- Total de atendimentos durante 10 anos
- Média de atendimentos durante 10 anos
- Últimos dois anos de atendimentos
- Gráfico representando atendimentos durante 10 anos

list_Ocupação.json

- Substituição dos valores no nulos por zero
- Soma do número de quartos/camas ocupados
- Especificações do hotel com três estrelas
- Descrição da tabela "Ocupação"

INSTITUTO POLITÉCNICO DE COIMBRA – ESCOLA SUPERIOR DE EDUCAÇÃO

CURSO LICENCIATURA DE COMUNICAÇÃO E DESIGN MULTIMÉDIA

Programação e Algoritmos II - Docente: João Orvalho

preços_1

Gráfico representando diferença de preços entre lazer e negócio

Gráfico global

Gráfico total de todos os dados

dataset-lugares-publicos-lugares-turisticos.csv

Renomeação das colunas das coordenadas

Eliminar linhas com dados em falta

Ordenar de forma decrescente todos os dados

Transformar os valores em strings

Transformar colunas em valores strings

Gráfico de coordenadas

INSTITUTO POLITÉCNICO DE COIMBRA – ESCOLA SUPERIOR DE EDUCAÇÃO

CURSO LICENCIATURA DE COMUNICAÇÃO E DESIGN MULTIMÉDIA

Programação e Algoritmos II - Docente: João Orvalho

Capítulo III - EstruturaOrganização repositório:

Dormidas nº estadia.json - contém dados relativos ao número de estadias turísticas no ano de 2019.

alojamento2019.csv - contém dados relativos a alojamentos em 2019 na localidade de Águeda.

Ano atendimento.csv - contém dados extraídos da Câmara Municipal de Águeda que nos fornece o número de atendimentos no posto Turismo do mesmo ao longo de 10 anos.

list_Ocupação.json - contém dados relativos a taxa de ocupação de 4 hotéis com diferentes categorias.

preços_1 - contém dados relativos a despesas turísticas tanto de lazer como de negócios, obtidas através de um inquérito realizado a 1393 hóspedes na cidade de Lisboa.

dataset-lugares-publicos-lugares-turisticos.csv - contém diversos dados de regiões turísticas do mundo, particularmente coordenadas.

Capítulo IV - API's usados

Esta base de dados utilizada neste relatório foi grande parte retirada de dados.gov.pt, no entanto utilizamos também dados fornecidos pela Câmara Municipal de Águeda.

Capítulo V - Dicionário de dados

Dormidas nº estadia.json

| Nome da coluna | Significado | Possíveis valores |
|------------------|---|-----------------------------|
| dormidas | Conjunto de dados de como a informação foi extraída, acerca dos dados de dormidas de 2019 | pandas.core.frame.DataFrame |
| dadosDormidas | Lista dos Dados de 2019 | pandas.core.frame.DataFrame |
| dados2019 | Conjunto de dados sobre as informações de dormidas no ano 2019 | pandas.core.frame.DataFrame |
| tamanho_tabela_1 | Tamanho tabela dados 2019 | Inteiro |
| dados | Renomeação das colunas | pandas.core.frame.DataFrame |
| dados_1 | Apresenta colunas importantes | pandas.core.frame.DataFrame |
| sem_valores | Total de dados em falta em dada coluna | pandas.core.series.Series |
| result_sc | Eliminar colunas com dados em falta | pandas.core.frame.DataFrame |
| rslt_dados2019 | Dados de dormidas em Portugal no ano de 2019 | pandas.core.frame.DataFrame |
| result_sl | Eliminar linhas com valores em falta | pandas.core.frame.DataFrame |

Alojamentos2019.csv

| Nome da coluna | Significado | Possíveis valores |
|------------------|--|-----------------------------|
| alojamento | Alojamentos de 2019 de Águeda | pandas.core.frame.DataFrame |
| al_sv | Soma de todos os valores em falta | Inteiro |
| al_sv_1 | Localização do dados em falta | Gráfico |
| ordem | Ordenar nomes dos alojamentos por ordem alfabética | pandas.core.frame.DataFrame |
| rslt_alojamentos | Moradias que se enquadram na categoria motel | pandas.core.frame.DataFrame |
| imgplot | Moradias que se enquadram na categoria motel | matplotlib.image.AxesImage |

INSTITUTO POLITÉCNICO DE COIMBRA – ESCOLA SUPERIOR DE EDUCAÇÃO
CURSO LICENCIATURA DE COMUNICAÇÃO E DESIGN MULTIMÉDIA

Programação e Algoritmos II - Docente: João Orvalho

Ano atendimento.csv

| Nome da coluna | Significado | Possíveis valores |
|----------------|--|-----------------------------|
| "line" | Visualização do ficheiro em formato de objeto | String |
| atendimentos | Atendimentos por ano no posto de Águeda | pandas.core.frame.DataFrame |
| at_10 | Soma do 10 anos de atendimentos | Inteiro |
| media_at | Média de atendimentos durante 10 anos | Inteiro |
| ultimos_anos | Últimos dois anos de atendimentos | pandas.core.frame.DataFrame |
| g_at | Gráfico representante dos anos de atendimentos | Gráfico |

list_Ocupação.json

| Nome da coluna | Significado | Possíveis valores |
|----------------|---|-----------------------------|
| ocupação | Conjunto de dados relativos a taxas de ocupação | pandas.core.frame.DataFrame |
| categoria | Substituição dos valores nulos por 0 | pandas.core.frame.DataFrame |
| soma_quartos | Soma dos quartos ocupados | Inteiro |
| soma_camars | Soma das camas ocupados | Inteiro |
| estrelas_3 | Especificações do hotel com 3 estrelas | pandas.core.series.Series |
| d_ocp | Descrição da tabela "ocupação" | pandas.core.frame.DataFrame |

preços_1

| Nome da coluna | Significado | Possíveis valores |
|----------------|---|-------------------|
| preços_1 | Conjunto de dados relativos aos preços | Lista |
| gra_preços | Diferença de preços entre Lazer e Negócio | Gráfico |
| global_1 | Diferença global de Lazer e Negócio | Gráfico |
| total | Total de todos os dados | Gráfico |

INSTITUTO POLITÉCNICO DE COIMBRA – ESCOLA SUPERIOR DE EDUCAÇÃO

CURSO LICENCIATURA DE COMUNICAÇÃO E DESIGN MULTIMÉDIA

Programação e Algoritmos II - Docente: João Orvalho

dataset-lugares-publicos-lugares-turisticos.csv

| Nome da coluna | Significado | Possíveis valores |
|----------------|--|-----------------------------|
| renomeação_at | Renomeação das colunas devido a serem nomes estrangeiros | pandas.core.frame.DataFrame |
| coordenadas | Colunas das coordenadas | pandas.core.frame.DataFrame |
| c_sl | Elimina as linhas com falta de dados nas colunas descritas | pandas.core.frame.DataFrame |
| descendente | Ordenar de forma descendente os dados | pandas.core.frame.DataFrame |
| string_value_1 | Transformar valores em strings | String |
| df | Transformar colunas em valores strings | pandas.core.frame.DataFrame |
| coordenada_1 | Transformar colunas em valores strings | Gráfico |

print(type(df))- Encontrar possíveis valores

Capítulo VI – Conclusão

Em suma podemos concluir que o presente projeto permitiu ao grupo aprender variadas informações sobre o turismo Português e Estrangeiro, mas acima de tudo, e o mais importante permitiu-nos aprender novos conteúdos programáticos e a cima de tudo implementar aqueles que foram aprendidos em aula.

Achamos deveras gratificante para a completação e finalização desta Unidade Curricular a execução deste projeto.

INSTITUTO POLITÉCNICO DE COIMBRA – ESCOLA SUPERIOR DE EDUCAÇÃO
CURSO LICENCIATURA DE COMUNICAÇÃO E DESIGN MULTIMÉDIA
Programação e Algoritmos II - Docente: João Orvalho

Parte III – Bibliografia

<https://dados.gov.pt/pt/>

<https://dados.gov.pt/pt/datasets/r/b9b406ab-bc80-448f-8c47-406f66b4e48a>

<https://dados.gov.pt/pt/datasets/r/029f7761-d427-404a-a639-a3eb1c83ffd8>

<https://ckan.sig.cm-agueda.pt/dataset/antendimentos-no-posto-de-turismo-de-agueda/resource/a2abb35b-62b1-4b9f-87b4-5a3a67558626>