

Uma Análise Quantitativa dos Meios de Hospedagem do Recife

Juliana Carvalho¹, Ângelo de Sant’Ana², Antônio Carlos³

*Centro de Informática - Universidade Federal de Pernambuco (UFPE)
Caixa Postal 7.851 – 50.732-970 – Recife – PE – Brasil
jco2@cin.ufpe.br, assd@cin.ufpe.br, acpr@cin.ufpe.br*

Abstract

O presente trabalho objetiva analisar a qualidade dos hotéis de acordo com as avaliações dos usuários do site Trivago na Grande Recife - PE. Para alcançar esse objetivo foi realizada uma análise qualitativa e exploratória dos dados. Também foi feita uma pesquisa bibliográfica de caráter exploratório sobre “turismo e hospedagens”, “qualidade e satisfação dos clientes” e “internet e os meios de hospedagens”. O artigo divide-se em cinco seções: introdução, fundamentação teórica, metodologia, análise dos dados e considerações finais. Pretende-se analisar diferentes pontos de um tema a fim de contribuir para a evolução da discussão do tema inovação em meios de hospedagem e fornecer novas bases para a produção de futuras pesquisas acadêmicas.

Keywords: Análise quantitativa, Turismo, Qualidade, Hospedagem, Algoritmos, Ciência dos dados

1. Introdução

Este projeto tem como objetivo colocar em prática os conhecimentos na área de Ciência dos Dados. Escolhemos uma base de dados para que após aplicadas as técnicas de Extração, Pré-processamento e Visualização pudéssemos
5 analisar e extrair informação dos dados. Também utilizamos de algoritmos de

¹Graduanda do curso Bacharelado em Sistemas de Informação da UFPE.

²Graduando do curso Bacharelado em Ciência da Computação da UFPE.

³Graduando do curso Bacharelado em Ciência da Computação da UFPE.

aprendizagem para que aumentar a qualidade da nossa análise. O trabalho está dividido em 5 seções. são apresentados o contexto que desencadeou a escolha do tema de pesquisa, objetivos e estrutura do trabalho aqui apresentado. A seção 2 chamado de Referencial Teórico, onde os principais conceitos e definições serão mostrados, para um real entendimento do trabalho. Na seção 3 deste documento será apresentado quais foram os métodos utilizados para chegar nos resultados desta análise. A seção 4 mostra os resultados da análise dos dados. Na seção 5, são descritas as principais contribuições.

2. Referencial Teórico

Esta seção contém uma síntese dos conceitos que foram estudados e que serviram para a execução da pesquisa.

3. Metodologia

Nesta seção serão explicados os métodos utilizados para o desenvolvimento deste trabalho.

3.1. Extração de dados

Neste projeto, para a realização da coleta dos dados, foi necessário a utilização do método de data-scraping . Nosso objetivo era recuperar as avaliações e informações de todos os hotéis da região da Grande Recife. O site de hospedagem escolhido foi o Trivago(<https://www.trivago.com.br/>). Antes de realizar a extração, todas as páginas foram baixadas para evitar problemas com mudanças no código fonte. Após a coleta, os foram salvos em um arquivo CSV(Comma-separated values), chamado "saidaFinalDataScraping".Para em seguida ser utilizado na etapa de pré-processamento.

3.2. Pré-processamento

Na etapa de pré-processamento, foram utilizadas diversas bibliotecas *Python* para realização desse processo. Com o auxilio dessas bibliotecas é possível fazer

manipulações e tratamento de vários conjuntos de dados.

O pré-processamento constituiu-se em:

- (1) Substituir os valores não informados por NaN;
- 35 (2) Remover as colunas que possuíam mais de 3 valores NaN;
- (3) Preencher os valores NaN com a média da coluna;
- (4) Remover as colunas;
- (5) Salvar dados tratados em um novo arquivo CSV nomeado "saidaPreProcessamento";

40 3.3. Técnica de Aprendizagem

Utilizamos regressão linear como técnica de aprendizagem, abiblioteca utilizada foi a *LinearRegression* e criamos um gráfico simples que mostra o resultado desta aplicação.

3.4. Visualização

- 45 Para a visualização foi utilizada a biblioteca *matplotlib*, criamos 2 gráficos simples para mostrar sua utilização.

4. Análise dos Resultados

5. Considerações Finais

References