

Apresentação

Aplicando o agrupamento de dados para identificar sobreposição de rotas

Jorge Nery * Iarley Moraes*

*Departamento de Ciência da Computação, Instituto de Computação, Universidade Federal da Bahia.

17 de agosto de 2024

Apresentação

Jorge Nery, Iarley Moraes

Introdução

Preparação e limpeza de dados

Metodologia

esultado

inal

xxxxXX - 2024.2 (1/1) 1/8

Introdução



O objetivo deste projeto é transformar as rotas de linhas de ônibus em grafos e avaliar a similaridade entre elas. A análise focará no particionamento e na concorrência das rotas, buscando identificar padrões que possam contribuir para a otimização do planejamento e da eficiência do sistema de transporte público

Base de Dados

Composta pelo rota das linhas operadas em 11/11/2019

Apresentação

Jorge Nery, Iarley Moraes

Introdução

Preparação e limpeza de dados

Metodologia

esultado

Final

Preparação e limpeza de dados



No processo de preparação para analisar os dados foram executadas algumas atividades de limpeza e organização dos dados

- 1. Removidas as informações de rotas duplicadas.
- 2. Garantida a interpretação das colunas numéricas.
- 3. Transformação da lista de pontos em Grafo
- 4. Plotagem de Gráficos para pre-visualização das informações
- 5. Redução dos dados

Apresentação

Jorge Nery, Iarley Moraes

Introdução

Preparação e limpeza de dados

Metodologia

Resultado

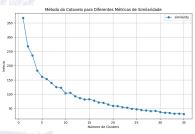
Final

Metodologia



Distancia e Similaridade

- Foram implementadas algumas formulas de calculo de distancia e similaridade.
- Foi criada uma matriz de similaridade
- Foi implementada uma função para localizar o numero de clusters usando o método cotovelo



Fonte: Gerada Usando MatPlotLib no JupyterNotebook

Apresentação

Jorge Nery, larley Morae

Introdução

Preparação e impeza de dados

Metodologia

Resultado

Einal

xxxxX - 2024.2 (1/2) 4/8

Metodologia



Clusters / Agrupamentos

- Detalhamento do Processo de Clusterização.
- Ferramentas e Software Utilizados
- Validação dos Resultados

Linha	Grupo
1225	4
1145	22
041001	21
114401	8
142001	25
150802	31
N035	15
1137	8
1652	17
0354	22
1148	2
1103	0

Tabela de linhas e seus respectivos grupos

Apresentação

Jorge Nery, larley Moraes

Introdução

reparação e mpeza de lados

Metodologia

Resultado

Einal

xxxxX - 2024.2 (2/2) 5/8

Resultados



A análise gerou uma matriz de similaridade que, juntamente com o método do cotovelo, possibilitou a identificação do número ideal de clusters. Esses clusters representam grupos de rotas com sobreposição significativa, permitindo uma visão mais clara de como as rotas se organizam e competem dentro do sistema de transporte

Desafios Enfrentados

- Volume de informações
- Tratamento dos dados
- Uso futuro e próximos passos



Fonte: Gerada Usando Florium no JupyterNotebook

Apresentação

Jorge Nery, Iarley Moraes

Introdução

Preparação e limpeza de dados

Metodologia

Resultado

Final

xxxX - 2024.2 (1/1) 6/8

Obrigado



Informações complementares

• Duvidas e Perguntas?

Códigos, Copia da Apresentação, Slides Manual de Execução dos Códigos: https://github.com/jorgenery/ufba-mestrado Apresentação

Jorge Nery, Iarley Moraes

Introdução

Preparação e limpeza de dados

Metodologia

Resultado

Final

xxxxX - 2024.2 (1/1) 7/8

Bibliografia e referências I





Apresentação

Jorge Nery, Iarley Moraes

Introdução

Preparação e limpeza de dados

Metodologia

Resultado

Final

xxxxX - 2024.2 (1/1) 8/8