

MC536 - Etapa 1 do Trabalho

Victor Palmerini 178061
Gabriel Francioli 172111

1) Resumo

O trabalho proposto tem como objetivo resumir, a partir do agrupamento e visualização, dados relacionados a diagnóstico e incidência de diabetes no Estados Unidos. Os dados são mapeados por estados e as análises estatísticas terão então esse objetivo de extrair informações relevantes sobre o dataset que será utilizado, como médias, desvios, agrupamento e identificação de padrões.

2) Lista de Requisitos

- Dados (porcentagens) são mapeados de acordo com o estado ou county
- Incidências são mapeadas por ano
- Incidências podem ocorrer de acordo com idade, gênero ou escolaridade, que são atributos relacionados a uma pessoa
- Para cada instância, precisamos saber o local e a porcentagem de incidências naquele local
- Locais são relacionados a um estado dos Estados Unidos
- Instâncias podem ser relacionadas a porcentagem de incidências, fatores de risco, práticas preventivas, fatores de risco por complicação ou estado terminal

3) Fonte dos Dados

Center of Disease Control and Prevention:

<https://gis.cdc.gov/grasp/diabetes/DiabetesAtlas.html#>

Neste site é possível encontrar dados demográficos relacionados à incidência e distribuição de diabetes pelos estados dos Estados Unidos. Há ainda filtros específicos para a busca de dados, como idade, gênero e escolaridade.

4) Modelagem Conceitual

