

Estatística multivariada na análise da saúde materna

Anna Caroline Pereira Costa
Santana Gonzalez

Profª Orientadora: Alessandra dos Santos

ENGENHARIA



XXIII ECCI

ENCONTRO
CIENTÍFICO CULTURAL
INTERINSTITUCIONAL

sisprime
cooperativa de crédito

INTRODUÇÃO

Com o avanço das tecnologias e o uso da Internet das Coisas (IoT), novos métodos de monitoramento de gestantes estão sendo adotados. Segundo Ahmed M., Kashem MA, Rahman M., Khatun S. (2020) em Bangladesh, um sistema baseado em IoT foi implementado para monitorar, em tempo real, as condições de saúde de mulheres grávidas em áreas rurais, permitindo a coleta contínua de dados vitais, como pressão arterial, glicemia, temperatura e frequência cardíaca.

Este estudo analisa uma base de dados com 1.014 registros de gestantes coletados por esse sistema, buscando identificar fatores de risco associados à saúde materna. A análise dos dados teve como objetivo auxiliar na compreensão das técnicas estatísticas discutidas em sala de aula.

METODOLOGIA

Os dados utilizados neste estudo foram coletados de gestantes em áreas rurais de Bangladesh, por meio de um sistema de monitoramento de saúde baseado na Internet das Coisas (IoT), disponível em: Conjunto de dados de risco à saúde materna.csv no <https://www.kaggle.com/datasets/csafrit2/maternal-health-risk-data/data>.

Variáveis do estudo: idade, pressão arterial, glicemia, frequência cardíaca, nível de risco (low, mid, high)

Técnicas utilizadas:

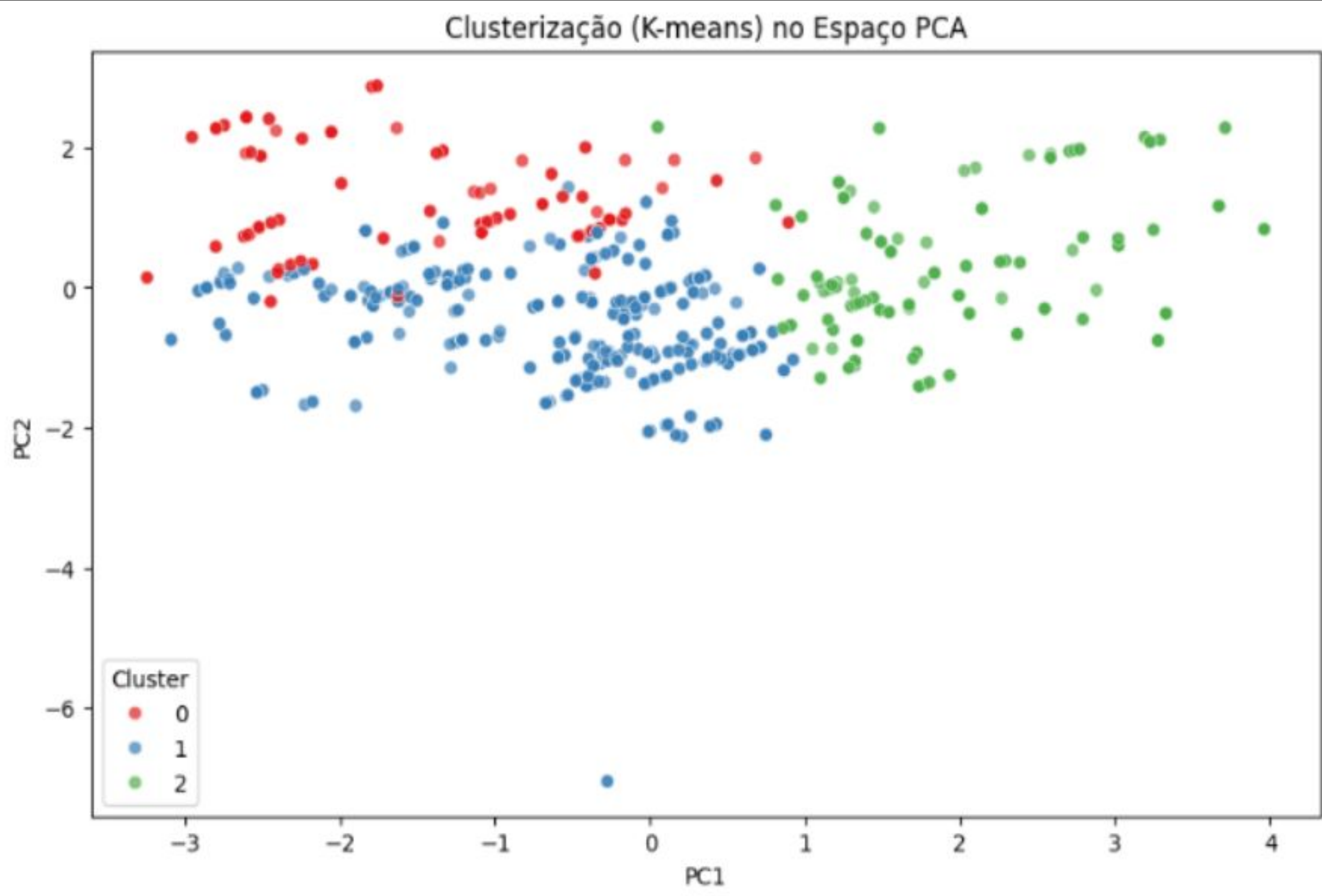
- ☐ Análise univariada: medidas descritivas (média, mediana, desvio padrão, mínimo e máximo) e gráficos.
- ☐ Análise bivariada: relação entre variáveis por meio de gráficos de dispersão e correlação
- ☐ Análise multivariada: Análise de Componentes Principais (PCA) para reduzir a dimensionalidade dos dados e identificar padrões de variabilidade entre as variáveis.
 - Análise foi realizada com Python, utilizando as bibliotecas Pandas, Seaborn e Matplotlib.

RESULTADOS

A análise dos dados obtidos aponta que em média as mulheres tinham 30,81 anos, sendo que a mais nova possuía 14 e a mais velha era de 70 anos. Com relação a pressão arterial, os resultados variaram de 120 mmHg (alta) e 60 a 80 mmHg (baixa).

No total, 40,46% apresentaram risco baixo, 33,14% risco médio e 26,82% risco alto.

Ao realizar a análise de componentes principais, a componente 1 explica 42% dos dados e a componente 2 20%. Totalizando 62% de variância explicada.



- Grupo 0: localizado majoritariamente na parte superior esquerda do gráfico. Pode representar um grupo de gestantes com características clínicas mais críticas, como pressão alta, idade mais avançada e glicemia elevada. Este grupo possivelmente se relaciona com o alto risco de complicações gestacionais.
- Grupo 1 espalhado de forma mais densa no centro do gráfico, agrupa gestantes com valores clínicos intermediários, ou seja, pacientes que não se destacam por extremos em nenhuma variável isolada. Pode estar relacionado ao risco médio, ou até mesmo conter uma mistura de pacientes com perfis pouco distintos entre si. A presença de sobreposição com os outros clusters sugere que este grupo representa uma zona de transição.
- Grupo 2 concentrado à direita do gráfico, agrupa gestantes com valores clínicos dentro da normalidade, como pressão arterial, glicemia e temperatura corporal em níveis saudáveis. Este cluster tem fortes indícios de representar o grupo de baixo risco.

A separação visual entre os clusters demonstra que existem padrões estruturais nos dados que permitem a identificação de perfis distintos.

CONSIDERAÇÕES FINAIS

- A utilização do PCA torna possível visualizar esses agrupamentos com clareza, facilitando a interpretação dos resultados.
- Abordagem evidencia que os dados têm características discriminatórias, o que reforça o valor clínico das variáveis coletadas.
- Limitações do estudo, como a ausência de variáveis adicionais (histórico médico e fatores sociais) e a natureza observacional dos dados, indicam a necessidade de pesquisas futuras mais abrangentes.

Em resumo, o estudo reforça a importância de estratégias inovadoras e acessíveis para monitorar e reduzir os riscos maternos em países em desenvolvimento.

REFERÊNCIAS

Artigo. Revisão e Análise de Fatores de Risco para a Saúde Materna em Áreas Remotas Utilizando a Internet das Coisas (IoT) - Ahmed M., Kashem MA, Rahman M., Khatun S. (2020) [https://www.kaggle.com/datasets/csafrit2/maternal-health-risk-dat a/data](https://www.kaggle.com/datasets/csafrit2/maternal-health-risk-data/data)