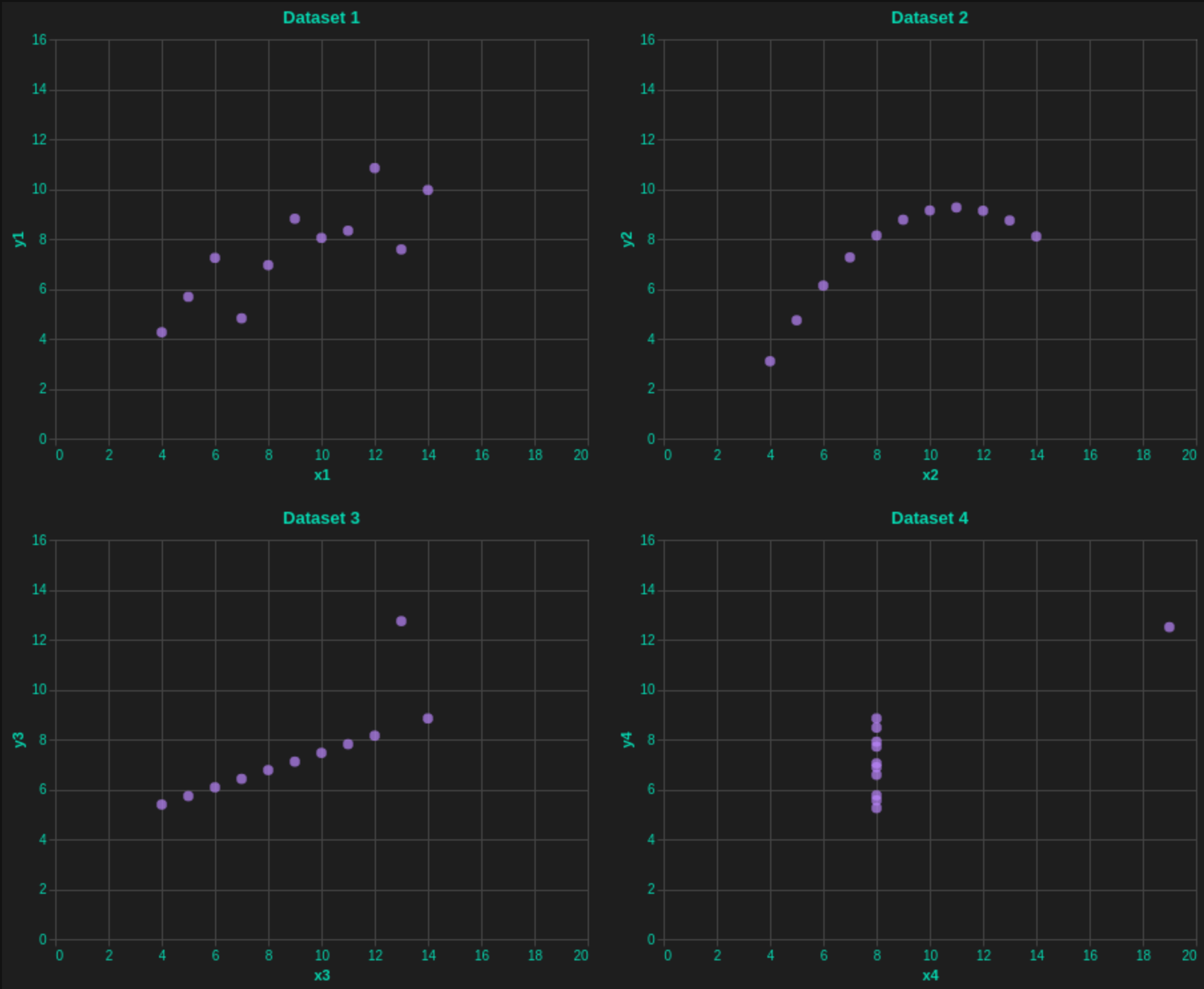


Quarteto de Anscombe

Introdução

O Quarteto de Anscombe consiste em quatro conjuntos de dados nos quais têm as mesmas estatísticas descritivas: média de x, média de y, variância de x, variância de y, correlação entre x e y, coeficiente de determinação e regressão linear. Porém, como veremos mais abaixo, os dados não são iguais. Esse é um famoso exemplo de como as estatísticas descritivas podem ser enganosas e por isso é importante visualizar os dados antes de tirar conclusões precipitadas. Tendo dito isso, para exibir os dados, utilizaremos a biblioteca Altair, que é uma biblioteca de visualização de dados baseada em Vega-Lite (a motivação para o uso dessa biblioteca se encontra descrita nas considerações finais). Por sua vez, segue abaixo os gráficos de dispersão com os conjuntos de dados do quarteto de Anscombe:



Considerações finais

Primeiramente, a intenção era utilizar a biblioteca Vega-Lite, mas ao entrar no site oficial da biblioteca é recomendado o uso da biblioteca Altair e, caso seja desejado utilizar o Vega-Lite, ainda assim é necessário importar o Altair, pois o Vega-Lite é encontrado como uma "sub biblioteca" da Altair no Python. A motivação para utilizar o Vega-lite surgiu desde outra disciplina, de Análise Exploratória de Dados e Visualização, onde ela foi mencionada como uma forma mais simplificada da biblioteca Vega (na qual é, de certa forma, "mais completa" e permite mais personalização, porém é muito complexa de usar). Além disso, era desejável tanto utilizar alguma biblioteca em Python (pela facilidade com a linguagem) quanto poder visualizar os dados em uma página web, o que demandaria ser necessário exportar o gráfico de alguma forma que fosse compatível com HTML. Dessa forma, a biblioteca Altair foi escolhida por unir essas duas características de maneira que conseguimos exportar nossas visualizações facilmente em diversos formatos, inclusive mais de um que nos permite usá-las em páginas web e, além disso, nos dá a possibilidade de implementar gráficos interativos de forma simplificada. Ao trabalhar com a biblioteca, foi possível perceber que ela parece implementar bem a "gramática de gráficos", de forma que vemos que de acordo com a sequência de modificação que fazemos na visualização, podemos ver uma sobreposição dos recursos gráficos, de forma que podemos utilizar isso a nosso favor para construir com mais facilidade gráficos mais complexos, com mais recursos, e coerentes. Por fim, a biblioteca é bem simples de usar e permite uma boa personalização dos gráficos e, principalmente, tem uma boa documentação, desde o básico para começar a trabalhar com a biblioteca de forma simples (inclusive permitindo implementar interatividade simples sem quase nenhum esforço), até recursos "mais avançados", que abrem a possibilidade de personalização dos gráficos de forma mais detalhada.

Observações

O código para gerar os gráficos e esta página HTML foram feitos em um notebook python, que pode ser encontrado em um documento separado. Tanto essa página HTML quanto o notebook python estão disponíveis no [repositório do GitHub](#).

Bibliografia

ALTAIR. Vega-Altair: Declarative Visualization in Python. Disponível em: [altair-viz.github.io/](#). Acessado em 27 de setembro de 2022.

WIKIPEDIA. Quarteto de Anscombe. Disponível em: [pt.wikipedia.org/](#). Acessado em 27 de setembro de 2022.

VANDERPLAS, Jake. Multiple Charts in one HTML. Disponível em: [github.com/](#). Acessado em 1 de outubro de 2022.

MATERIAL. Dark theme. Disponível em: [material.io/](#). Acessado em 2 de outubro de 2022.

ADOBE. Adobe Color. Disponível em: [color.adobe.com/pt/create/](#). Acessado em 2 de outubro de 2022.

ALTAIR. Top-Level Chart Configuration. Disponível em: [altair-viz.github.io/](#). Acessado em 2 de outubro de 2022.