**DSDV**

**1.a Chuong:**What is its data, types of data

* Percentage of methylation: Ratio
* Cancer types: Nominal
* Organs: Nominal

**b) Column charts encode two attributes using a line mark with the vertical spatial position channel for the quantitative attribute, and the horizontal spatial position channel for the categorical attribute.**

How the author use marks and channels to encode data

* Channel
* Height: to describe the percentage of methylation
* Color: to represent the genetic mutation that causes cancer.
* Position: to describe the genetic mutation (x-axis) and the cancer-containing organ (z-axis)
* Mark: 3D mark - Columns / Volume to display each data point entry\*\*

c) Describe problems:

* 3-Dimension là sai nha nha
* Lack of coordinate axis names (Ox, Oy) (Thiếu label data cho các trục)
* Color ko có ý nghĩa (đã đc xác định r OOy)
* Few of the bars are hidden behind the other bars

One of the golden rules of data visualization goes like this: never use 3D. Repeat after me: never use 3D. The only exception is if you are actually plotting a third dimension (and even then, things get really tricky really quickly, so take care when doing this)—and you should never use 3D to plot a single dimension. As we saw in the pie chart example previously, 3D skews our numbers, making them difficult or impossible to interpret or compare.

Adding 3D to graphs introduces unnecessary chart elements like side and floor panels. Even worse than these distractions, graphing applications do some pretty strange things when it comes to plotting values in 3D. For example, in a 3D bar chart, you might think that your graphing application plots the front of the bar or perhaps the back of the bar. Unfortunately, it’s often even less straightforward than that.

3D volume for bars makes it harder to compare methylation percentage

d) redesign

* Đổi sang 2 chart 2D:  
  (1) Cancer type (x-axis) and the cancer-containing organ (z-axis)  
  (2) Cancer type (x-axis) and percentage of methylation (y-axis)
* Use Identity Channels: Categorical Attributes to represent the genetic mutation that causes cancer: Color hue

\*\*idk câu d

question2:

import matplotlib.pyplot as plt

GPA = [80,70,90,100,60,70,90,90,50,60,85,90,70,70,80,90,100,70,90,60]

GroupID = [1,1,1,1,2,2,2,2,3,3,3,4,4,4,5,5,5,6,6,6]

major = ['IT', 'CS', 'CS', 'IT', 'DS', 'DS', 'DS', 'CS', 'CS', 'IT', 'CS', 'DS', 'DS', 'DS', 'CS', 'IT', 'IT', 'DS', 'CS', 'CS']

Seat = [2,3,4,12,15,16,25,26,9,10,11,31,32,39,5,6,7,34,35,36]

fig, ax = plt.subplots()

for i in range(len(GPA)):

ax.scatter(Seat[i], GPA[i], c='b')

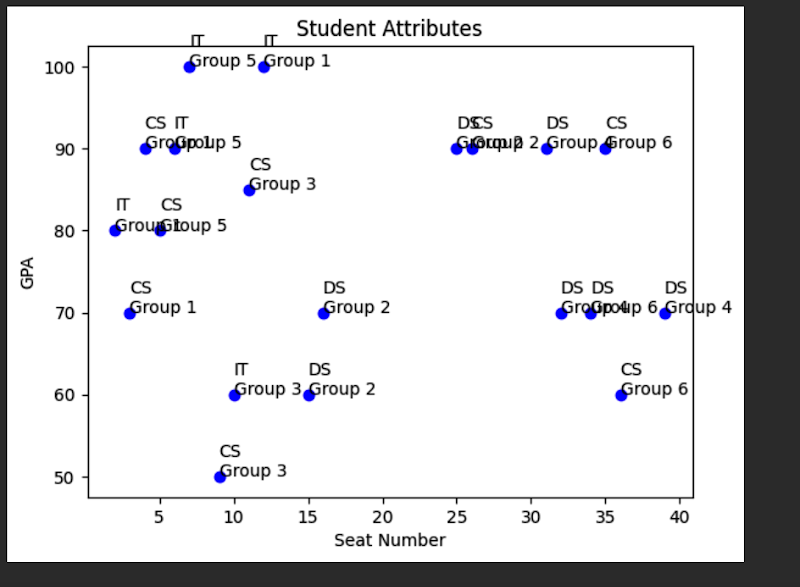
ax.text(Seat[i], GPA[i], f'{major[i]}\nGroup {GroupID[i]}')

ax.set\_xlabel('Seat Number')

ax.set\_ylabel('GPA')

ax.set\_title('Student Attributes')

plt.show()



Student in a class

- GPA: [Magnitude Channels: Ordered Attributes] Color Saturation

- GroupID (6 class): [Identity Channels: Categorical Attributes] Color Hue

Hai attribute này đều có thể dùng màu để thể hiện. Tụi m chọn một trong hai

- Major (3 class): [Identity Channels: Categorical Attributes] Shape

- Seat: Position (as in Figure 2)

* Link for girouD: Manipulate view. CHọn một seat -> hiện những bạn cùng nhóm -> ẩn những bạn khác nhóm

Allowing users to select one or more elements of interest in a vis is a fundamental action that supports nearly every interactive idiom. The output of a selection operation is often the input to a subsequent operation. In particular, the change choice is usually dependent on a previous select result. (Đọc trong sách thêm)

import numpy as np

import matplotlib.pyplot as plt

GPA = [80,70,90,100,60,70,90,90,50,60,85,90,70,70,80,90,100,70,90,60]

GroupID = [1,1,1,1,2,2,2,2,3,3,3,4,4,4,5,5,5,6,6,6]

major = ['IT', 'CS', 'CS', 'IT', 'DS', 'DS', 'DS', 'CS', 'CS', 'IT', 'CS', 'DS', 'DS', 'DS', 'CS', 'IT', 'IT', 'DS', 'CS', 'CS']

Seat = [2,3,4,12,15,16,25,26,9,10,11,31,32,39,5,6,7,34,35,36]

fig, axs = plt.subplots(nrows=3, ncols=2, figsize=(10, 10))

for group\_id in set(GroupID):

group\_members = [seat for seat, group in zip(Seat, GroupID) if group == group\_id]

row = (group\_id - 1) // 2

col = (group\_id - 1) % 2

ax = axs[row][col]

ax.set\_xticks(np.arange(1, 7))

ax.set\_yticks(np.arange(1, 8))

ax.set\_xticklabels(np.arange(1, 7))

ax.set\_yticklabels(np.arange(1, 8))

ax.set\_title(f'Group {group\_id}')

for member\_seat in group\_members:

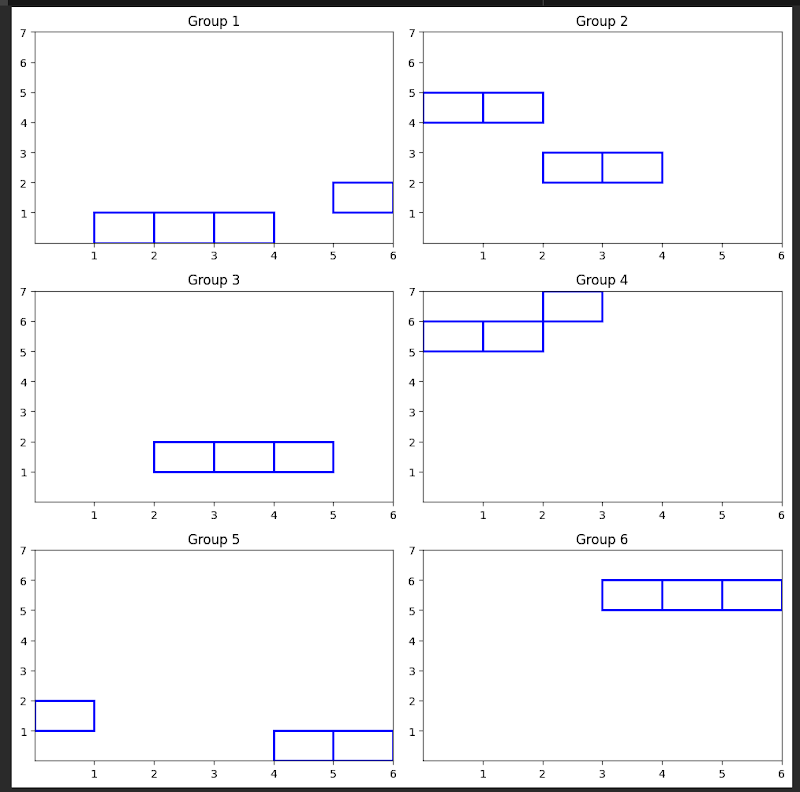
row = (member\_seat - 1) // 6

col = (member\_seat - 1) % 6

ax.add\_patch(plt.Rectangle((col, row), 1, 1, fill=False, edgecolor='b', lw=2))

plt.tight\_layout()

plt.show()



câu 3 : 