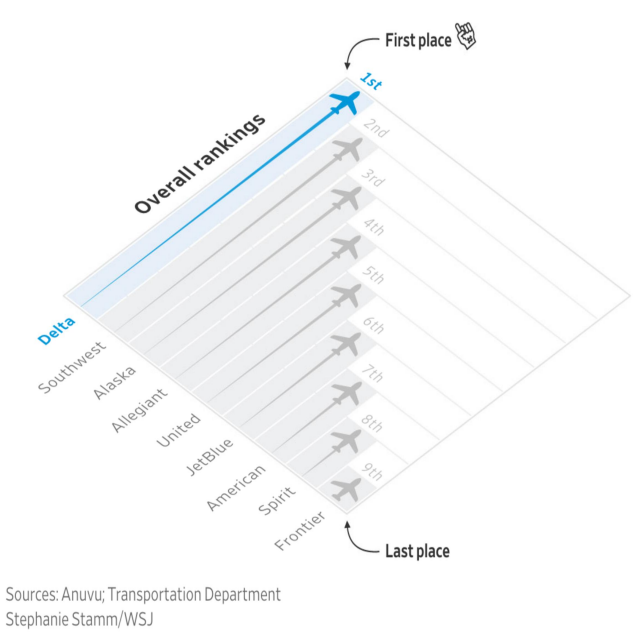
Homework 2

Github Repo: https://github.com/Rongxuan-Zhou/EECE5642-Data-Visualization/tree/main/HW2

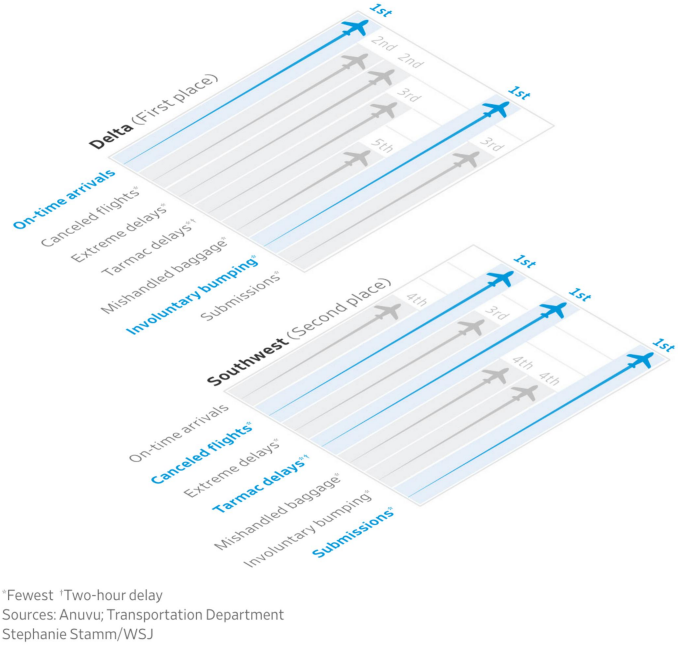
1. Visualization Design

a. Bad Design:

https://junkcharts.typepad.com/junk_charts/2025/01/ranks-labels-metrics-data-and-alignment.html, contained graphs are attached below:



junk chart1



junk chart2

Data-ink Ratio Issues:

The 45-degree slanted design creates excessive non-essential white space;
The airplane icons are decorative and convey no actual information;
Redundant labels and legends occupy too much space.

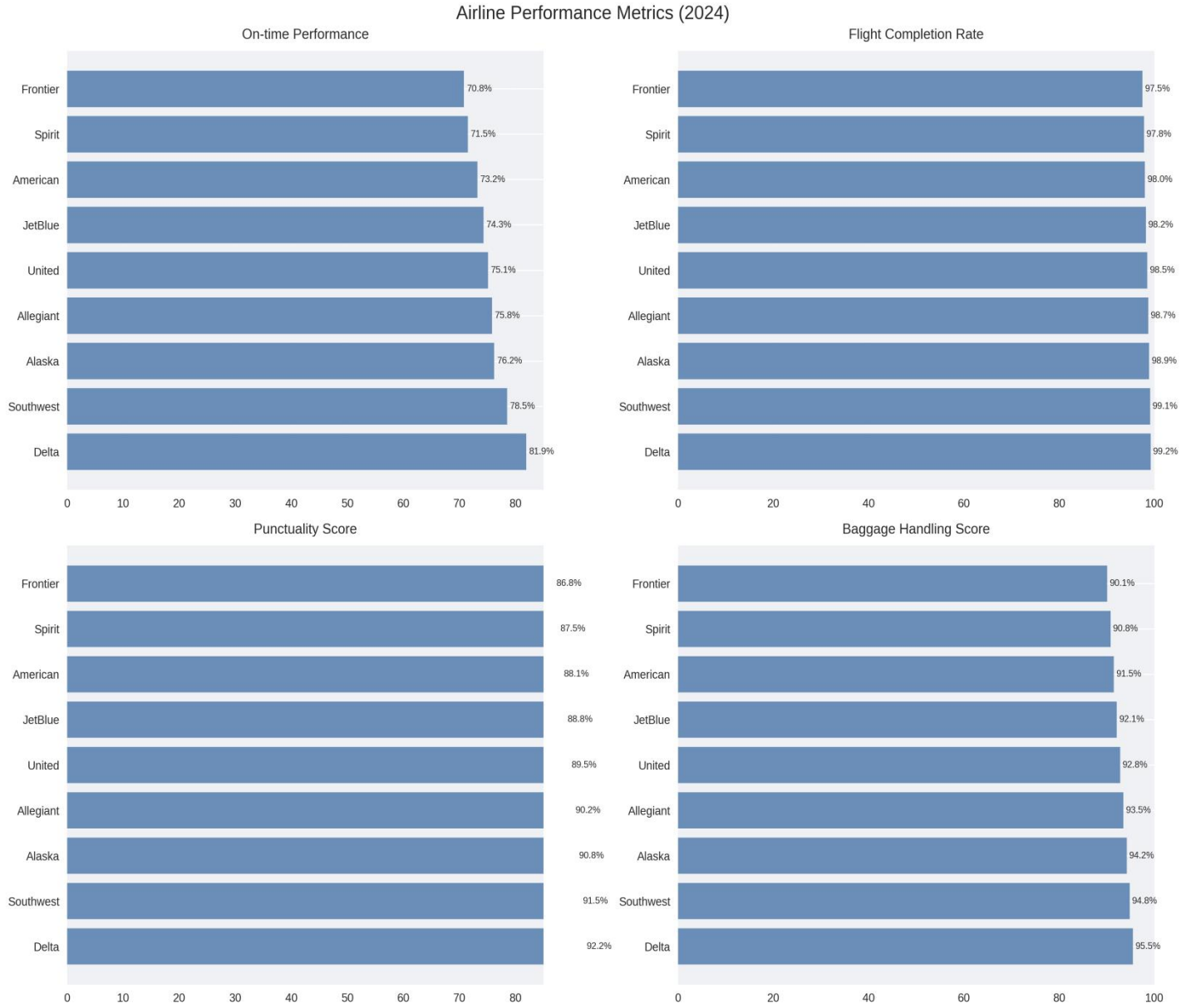
Lie Factor Analysis:

The reverse ranking representation can be misleading;
Asterisk annotations are needed to explain meanings, increasing cognitive load;
Using longer bars to represent "best" contradicts intuitive understanding.

Clarity Issues:

Text labels at 45-degree angles are difficult to read;
Multi-level information display complicates simple data;
Requires head tilting for proper information reading.

b. New Design:



Readability Improvements:

Use horizontal bar charts to avoid tilted text;
Display performance values directly instead of rankings;
Use clear metric names and avoid negative phrasing.

Information Display Optimization:

Group related metrics together;
Use gridlines to assist value comparison;
Add percentage labels for intuitive performance display.

Visual Enhancement:

Remove decorative elements (such as airplane icons);
Use subtle gridlines;
Adopt a clear color scheme.

Functional Improvements:

Display actual performance data directly;
Simplify information hierarchy;
Provide intuitive data comparison.

2. Color

a. Coding Resource:

The colorsys module from Python's standard library;

Matplotlib library for creating color swatches;

NumPy library for numerical computations;

Colormath library for more precise color space conversions;

EasyRGB's color conversion formulas (http://www.easyrgb.com/en/math.php);

Wikipedia's color space conversion entries;

Bruce Lindbloom's color transformation equations (http://www.brucelindbloom.com).

b. Results:

R = 137/255 = 0.5373

G = 56/255 = 0.2196

B = 146/255 = 0.5725

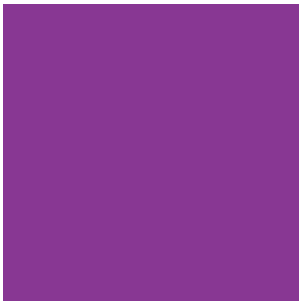
XYZ: (0.4034, 0.3126, 0.5808)

xyY: (0.3111, 0.2411, 0.3126)

CMYK: (0.0616, 0.6164, 0.0000, 0.4275)

HSV: (0.8167, 0.6164, 0.5725)

HSL: (0.8167, 0.4455, 0.3961)



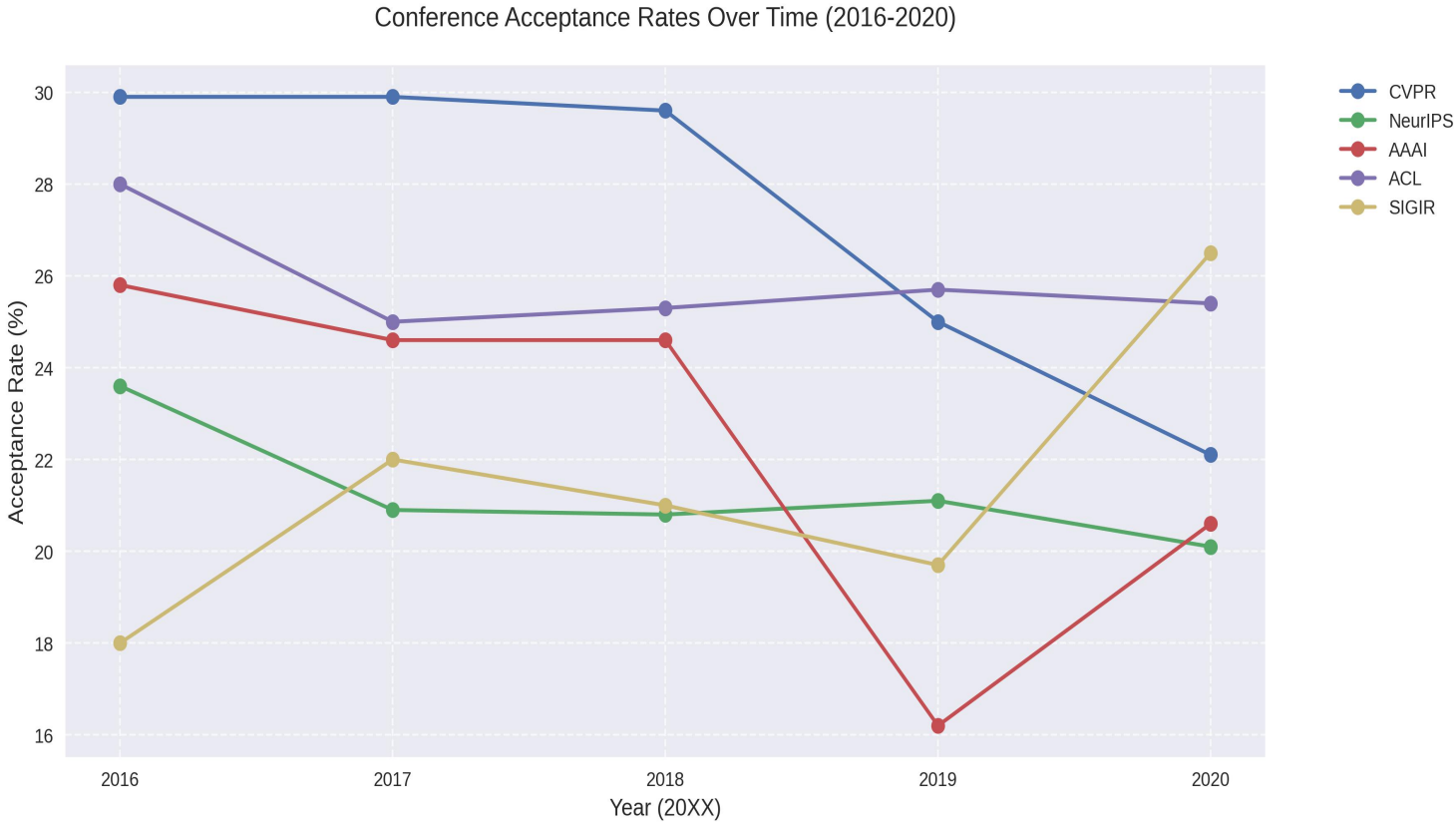
This color is a purple hue, leaning towards magenta in the RGB color space, with medium brightness and moderate saturation.

3. Table & Graph

In this section, partial dataset ranging from year 2016 to 2020 is picked, with generated table and graph attached below:

Conference Acceptance Rates - Table View (2016-2020)

Conference (Name'Year)	Acceptance rate	Num. of accepted papers	Num. of total submissions
CVPR'16	29.9%	643	2145
CVPR'17	29.9%	783	2620
CVPR'18	29.6%	979	3303
CVPR'19	25.0%	1294	5160
CVPR'20	22.1%	1470	6656
NeurIPS'16	23.6%	549	2403
NeurIPS'17	20.9%	678	3240
NeurIPS'18	20.8%	1011	4856
NeurIPS'19	21.1%	1428	6743
NeurIPS'20	20.1%	1900	9454
AAAI'16	25.8%	549	2132
AAAI'17	24.6%	638	2590
AAAI'18	24.6%	933	3800
AAAI'19	16.2%	1150	7095
AAAI'20	20.6%	1591	7737
ACL'16	28.0%	231	825
ACL'17	25.0%	195	751
ACL'18	25.3%	258	1018
ACL'19	25.7%	447	1737
ACL'20	25.4%	571	2244
SIGIR'16	18.0%	62	341
SIGIR'17	22.0%	78	362
SIGIR'18	21.0%	86	409
SIGIR'19	19.7%	84	426
SIGIR'20	26.5%	147	555



Visualization Comparison Analysis:

a. Table

Pros: Precise presentation of exact values;
Easy to look up specific numbers;
Good for comparing individual values;
Shows all three metrics (acceptance rate, accepted papers, total submissions).

Cons: Difficult to see trends over time;
Takes more space to display;
Requires more time to process information.

b. Graph

Pros: Clear visualization of trends over time;
Easy to compare patterns between conferences;
Intuitive understanding of acceptance rate changes;
Compact representation of temporal patterns.

Cons: Only shows acceptance rate (not paper counts);
Can become cluttered with too many conferences;
Less precise for exact values.

4. Visual Perception and Cognition

Analyze the two pics from the perspective of visual perception and cognitive psychology.

a. Visual Perception Level

Physical Similarity: Both images use the same bright yellow (#FFFF00) background;
Three characters arranged with equal spacing, balanced visual weight;
Using the same geometric sans-serif font with consistent stroke width.

Gestalt Principles:

Principle of Proximity: Three characters are perceived as a unit due to close spacing
Principle of Continuity: Character arrangement creates a sense of sequence
Principle of Closure: Yellow background encloses the character combination into a visual unit

b. Cognitive Processing Level

Context Effects:
Left image: "B" placed between A and C, alphabet sequence knowledge activated with the alphabet mental model (A→B→C)
Right image: "B" placed between 12 and 14, mathematical pattern knowledge activated (right image: arithmetic sequence, difference of 1) with the numerical sequence model (12→13→14)

Cognitive Processing:
Automatic processing: Quick recognition of basic letter and number shapes
Controlled processing: Inferring different meanings of "B" based on context
Left image "B" → Letter "B"
Right image "B" → Substitute symbol for number "13"

c. Cognitive Conflict and Resolution

Cognitive Conflict:
Symbol ambiguity: Same "B" symbol creates cognitive competition in different contexts
Category switching: Need to switch between letter and number systems

Resolution Mechanisms:
Context dependency: Using surrounding information to resolve ambiguity
Experience application: Applying existing sequence knowledge (alphabet order/numerical progression)
Pattern completion: Automatically filling in missing sequence items