# Non-Research Tips for Information Science Researchers (Summer 2024)

#### May 1, 2024 Week 4: Tables and plots

https://non-research-tips.github.io/2024



Yusuke Matsui (UTokyo)

### Schedule

Date (2024)	Contents	Presented by
Week 1, Apr 10	Introduction. Review of fundamental concepts	<del>Yusuke, Koya, Yuki, Jun</del>
Week 2, Apr 17	Equations and pseudo-codes	<del>Yusuke Matsui</del>
Week 3, Apr 24	Presentation	Koya Narumi
Week 4, May 1	Tables and plots	Yusuke Matsui
Week 5, May 8	Figures	Koya Narumi
Week 6, May 22	Videos	Koya Narumi
Week 7, May 29	Invited Talk 1	Dr. Yoshiaki Bando (AIST)
Week 8, June 5	Invited Talk 2	Prof. Katie Seaborn (Tokyo Tech)
Week 9, June 12	GitHub in depth	Yusuke Matsui
Week 10, June 19	Automation of research and research dissemination (Web, Cloud, Cl/CD)	Jun Kato
Week 11, June 26	Research community	Jun Kato
Week 12, July 3	3DCG illustrations	Yuki Koyama
Week 13, July 10	Final presentations	_





Name	Height (cm)	Weight (kg)	
LeBron James	206	113	
Anthony Davis	208	115	1
	The table tables in	e looks bet top-confe	ter! I often find this kinds of rence papers.



















EfficientNet

DenseNet

AlexNet

VGG

ResNet

Architecture

that DenseNet's scores are scattered.



## Reference

#### booktabs

- S. Fear, "Publication quality tables in LaTeX", Official documentation for booktabs in CTAN, 2020. <u>http://mirrors.ctan.org/macros/latex/contrib/booktabs/booktabs.pdf</u>
- M. Püschel, "Small Guide to Making Nice Tables", <u>https://people.inf.ethz.ch/markusp/teaching/guides/guide-tables.pdf</u> Introduction of booktabs (I was inspired by this document.)

#### **Boxplot**

M. Streit and N. Gehlenborg, "Bar charts and box plots", Nature Methods, 2014. <u>https://www.nature.com/articles/nmeth.2807</u>

#### The original document for this lecture

▶ 松井勇佑, "とにかくbooktabsを使おう", GitHub, 2022. <u>https://github.com/mti-lab/use\_booktabs\_anyway</u>



- 公式ドキュメント。細かいオプションなどは CTAN の公式ドキュメントを参照しましょう。
- Markus Püschel, "Small Guide to Making Nice Tables". 本資料と同じ立ち位置の、booktabs 紹介記事です。

本文章は上記文献と松井の経験則をペースにしています。間違いや、より良い方法もあると思います。提案やコメントやミスの指摘など

#### I'd like to translate it into English, but I have no time... PR welcome!

# Powerpoint?

- > The contents of today's lecture are for rendering tables by TeX.
- When you want to render tables in Powerpoint, you can imitate the TeX's one, like:

	Name	Height (cm)	Weight (kg)
	LeBron James	206	113
But I don't like	Anthony Davis	208	115
this space			

> Or you can do it in other ways, like:

Example	Meaning
$a \in [2, 7]$	$2 \le a \le 7$ More informative than $a \in \mathbb{R}$
$a \in (2,7)$	2 < a < 7
$a \in [2,7)$	$2 \le a < 7$
$a \in \{2, 7\}$	a = 2  or  a = 7
$a \in \{2,, 7\}$	If naturally interpreted, $a = 2$ or $a = 3$ or or $a = 7$ .

# Disclaimer

- > The contents of today's lecture are based on Matsui's rule of thumb.
- Please consider the contents as guidelines and apply them to your field at your discretion.
- > If you have better tips, please always let me know!
- I would like to offer you helpful tips and enhance the quality of your papers. (All of you!)

Masked Reviewer ID: A Review: Question Please briefly describe the paper's contributions, and list its positive and negative points. Overall Rating	The authors start from a small (16 centroids) local descriptor x by descending a tree wher to select the subtree. A posting list (aka inv +: an alternalive to the hierarchical k-means -: bad method, flawed experiments Definitely Reject			

Can we systematically overcome the typical mistakes that beginners make?

I have been thinking about such a thing for more than a decade ...

The review for my first CVPR submission (2013)

### **Tables**

- Basics
- Row-oriented structure
- Row grouping
- Row hierarchization
- Column hierarchization
- Partial horizontal line (cmidrule)
- Column to row
- Flowchart

# <u>Plots</u>

- Basics
- Bar chart to box plot
- Control parameter
- > Misc

## **Tables**

# Basics

- Row-oriented structure
- Row grouping
- Row hierarchization
- Column hierarchization
- Partial horizontal line (cmidrule)
- Column to row
- Flowchart

# <u>Plots</u>

- Basics
- Bar chart to box plot
- Control parameter
- > Misc

#### > Follow the 5 steps below.

Name	Height (cm)	Weight (kg)
LeBron James	206	113
Anthony Davis	208	115

```
\begin{tabular}{|c|c|c|} \hline
Name & Height (cm) & Weight (kg) \\ \hline
LeBron James & 206 & 113 \\ \hline
Anthony Davis & 208 & 115 \\ \hline
\end{tabular}
```

```
NameHeight (cm)Weight (kg)LeBron James206113Anthony Davis208115
```

#### \usepackage{booktabs}

```
\begin{tabular}{@{}lll@{}} \toprule
   Name & Height (cm) & Weight (kg) \\ \midrule
   LeBron James & 206 & 113 \\
   Anthony Davis & 208 & 115 \\ \bottomrule
  \end{tabular}
```

#### $\succ$ Follow the 5 steps below.

Name	Height (cm)	Weight (kg)	<pre>\begin{tabular}{ c c c } \hline Name &amp; Height (cm) &amp; Weight (kg) \\ \hline</pre>
LeBron James	206	113	LeBron James & 206 & 113 \\ \hline
Anthony Davis	208	115	Anthony Davis & 208 & 115 \\ \hline
Nama	Height (em)	Weight (ltg)	Step 1: import the package.
	meight (cm)	weight (kg)	
LeBron James	206	113	<pre>\begin{tabular}{@{}lll@{}} \toprule Name &amp; Height (cm) &amp; Height (kg) )) &gt; midpule</pre>
Anthony Davis	208	115	LeBron James & 206 & 113 \\ Anthony Davis & 208 & 115 \\ \bottomrule \end{tabular}

#### Step 2: Delete all vertical lines. $\succ$ i.e., { | c | c | c | } $\rightarrow$ {ccc} $\succ$ If you table is well structured, you don't need vertical lines. > If you think you need vertical lines, I recommend splitting the table into smaller ones \begin{tabular}{ \hline Weight (kg) Height (cm) Name Name & Height (cm) & Weight (kg) \\ \hline ♣ 206 & 113 \\ \hline LeBron James 206113LeBron James Anthony Davis & 208 & 115 \\ \hline Anthony Davis 115208\end{tabular} \usepackage{booktab Height (cm) Weight (kg) Name \begin{tabular}{@{}lll@{}} \toprule LeBron James 206113Name & Height (cm) & Weight (kg) \\ \midrule Anthony Davis 208115LeBron James & 206 & 113 \\ Anthony Davis & 208 & 115 \\ \bottomrule \end{tabular}



#### $\succ$ Follow the 5 steps below.

Name	Height (cm)	Weight (kg)	<pre>\begin{tabular}{ c c c } \hline Name &amp; Height (cm) &amp; Weight (kg) \\ \hline</pre>
LeBron James Anthony Davis	206 208	113 115	Step 4: Put a magical spacer symbols, " $@{}$ " > i.e., {111} $\Rightarrow$ { $@{}$ 111 $@{}$ }
			<ul> <li>This eliminates excess spaces.</li> <li>Many papers forget this.</li> </ul>
Name	Height (cm)	Weight (kg)	\usepackage{booktabs}
LeBron James Anthony Davis	206 208	113 115	<pre>\begin{tabular}{@{}]1@{}} \toprule Name &amp; Height (cm) &amp; Weight (kg) \\ \midrul LeBron James &amp; 206 &amp; 113 \\</pre>
			Anthony Davis & 208 & 115 \\ \bottomrule \end{tabular}

20

/
---

w/ "@{}"

Name	Height (cm)	Weight $(kg)$
LeBron James	206	113
Anthony Davis	208	115

Name	Height (cm)	Weight (kg)	
LeBron James	206	113	
Anthony Davis	208	115	

 Weird spaces... 
 Remember "invisible lines" in Week 3!



Step 5: Use " → i.e., ✓ The top ✓ The ne ✓ The bo ✓ Other → Well-struc	<pre>top/mid/bo top/mid/bo top/m</pre>	ttomrule" \toprule \midrule e → \bottom Delete! require horize	rule ontal lines just a bit.
Anthony Davis	208	115	y Davis & 208 & 115 \\ hline
Name	Height (cm)	Weight (kg)	\usepackage{booktabs}
LeBron James	206	113	\begin{tabular}{@{}lll@{}} (toprule)
Anthony Davis 208	208	208 115	LeBron James & 206 & 113 \\
			Anthony Davis & 208 & 115 (bottomrule)
			22

#### > Follow the 5 steps below.



You can always apply these 5 steps, and your table will be much more beautiful!

Name	Height (cm)	Weight (kg)
LeBron James	206	113
Anthony Davis	208	115

```
\begin{tabular}{|c|c|c|} \hline
Name & Height (cm) & Weight (kg) \\ \hline
LeBron James & 206 & 113 \\ \hline
Anthony Davis & 208 & 115 \\ \hline
\end{tabular}
```



#### \usepackage{booktabs}

```
\begin{tabular}{@{}lll@{}} \toprule
   Name & Height (cm) & Weight (kg) \\ \midrule
   LeBron James & 206 & 113 \\
   Anthony Davis & 208 & 115 \\ \bottomrule
  \end{tabular}
```

### **Tables**

- Basics
- Row-oriented structure
- Row grouping
- Row hierarchization
- Column hierarchization
- Partial horizontal line (cmidrule)
- Column to row
- Flowchart

# **Plots**

- Basics
- Bar chart to box plot
- Control parameter
- > Misc

- > A table should be **a stack of rows**, i.e., **row-oriented**.
- The best structure: "header" + "row1" + "row2" + ...

Header	Name	Height (cm)	Weight (kg)	> Each row should be
Row1	LeBron James	206	113	$\succ$ i.e., instances (rows) of
Row2	Anthony Davis	208	115	the same class (header)

> Each row should be semantically interchangeable.

Name	Height (cm)	Weight (kg)
Anthony Davis	208	115
LeBron James	206	113

 Usually, rows are sorted by some criteria.
 Yet, rows should be interchangeable.

- > A table should be a stack of rows, i.e., row-oriented.
- The best structure: "header" + "row1" + "row2" + ...



- For each column,
  - ✓ The header defines the type, i.e., "int".
  - ✓ Each row shows the value, i.e., "34".
- > Seems obvious? But it's not easy to strictly follow this principal.



- From S. Ren+, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", TPAMI 2017 (30K+ citation paper)
- Object detection papers do not follow this principal (because #class is ~20 and can fit the paper, thus it's possible to write everything one line).
- > One may restructure this table clearer as we'll show later, but it requires more spaces.
- > Everything up to the situation!

	рп	псіраі.																			,	,	,
	-	-					Т	уре	2														
																1	mbike	person	plant	sheep	sofa	train	tv
		3.7						bottle	bus	car	cat	chair	COW	table	dog	horse	mone.	(0.0	30.1	65.4	70.2	75.8	65.8
	# box	data	mAP	areo	bike	bird	boat	Dotte	77.2	78.2	82.0	40.7	72.7	67.9	79.6	79.2	73.0	69.0	31.8	70.1	74.8	80.4	70.4
method	# 004	07	66.9	74.5	78.3	69.2	53.2	36.6	11.5	78.6	86.7	42.8	78.8	68.9	84.7	82.0	76.6	69.9	27.2	71.4	62.5	77.4	66.4
SS	2000	07,12	70.0	77.0	78.1	69.3	59.4	38.3	81.6	70.0	78.0	50.7	78.0	61.1	79.1	81.9	72.2	75.9	37.2	(0.2	67.3	81.1	67.6
SS	2000	0/+12	68.5	74.1	77.2	67.7	53.9	51.0	75.1	79.2	/0.9	52.2	75.3	67.2	80.3	79.8	75.0	76.3	39.1	68.5	72.0	83.0	72.6
RPN*	300	07	(0.0	70.0	80.6	70.1	57.3	49.9	78.2	80.4	82.0	52.2	81.9	65.7	84.8	84.6	77.5	76.7	38.8	73.6	73.9	96.6	78.9
RPN	300	07	69.9	76.5	79.0	70.9	65.5	52.1	83.1	84.7	86.4	52.0	01.2	70.8	85.9	87.6	80.1	82.3	53.6	80.4	75.8	00.0	1012
RPN	300	07+12	73.2	70.0	\$2.0	77.7	68.9	65.7	88.1	88.4	88.9	63.6	86.5	70.0									
RPN	300	COCO+07+12	78.8	84.3	02.0	111			_														

# Exceptional cases: e.g., average values

- > There are several exceptions, such as a row of average values
- Such row is a special and cannot be swapped to other rows (usually placed at the bottom).

Name	Height (cm)	Weight (kg)	Name	Height (cm)	Weight (k
LeBron James	206	113	LeBron James	206	113
Anthony Davis	208	115	Anthony Davis	208	115
D'Angelo Russell	193	88	D'Angelo Russell	193	88
Average	202	105.3	Average	202	105.3

Creating tables: Following the row-oriented structure as much as possible while aiming for the clearest presentation.

### **Tables**

- Basics
- Row-oriented structure
- Row grouping
- Row hierarchization
- Column hierarchization
- Partial horizontal line (cmidrule)
- Column to row
- Flowchart

# <u>Plots</u>

- Basics
- Bar chart to box plot
- Control parameter
- > Misc

> You can group rows to improve the readability.

#### A good table. Rows are interchangeable.

Method	k	Runtime (ms)	Accuracy	Method	k	Runtime (ms)	Accuracy	Method	k	Runtime (ms)	Accuracy
XYZ	16	0.32	0.21	XYZ	16	0.32	0.21	XVZ.	16	0.32	0.21
XYZ	32	0.61	0.44		32	0.61	0.44	XYZ	32	0.61	0.44
Ours	9	0.47	0.26	Ours	9	0.47	0.26		9	0.47	0.26
Ours	18	0.99	0.77		18	0.99	0.77	Ours	18	0.99	0.77

However, there are repetitive descriptions, which makes the table unnecessarily wordy.

> You can group rows to improve the readability.

Method	k	Runtime (ms)	Accuracy	Method	k	Runtime (ms)	Accuracy	Method	k	Runtime (ms)	Accuracy
XYZ	16	0.32	0.21	XYZ	16	0.32	0.21	XVZ	16	0.32	0.21
XYZ	32	0.61	0.44		32	0.61	0.44	AIL	32	0.61	0.44
Ours	9	0.47	0.26	Ours	9	0.47	0.26		9	0.47	0.26
Ours	18	0.99	0.77		18	0.99	0.77	Ours	18	0.99	0.77

You can simply group some rows.
 Delete unnecessary descriptions.

> You can group rows to improve the readability.

#### > You can further draw a \midrule

Method	k	Runtime (ms)	Accuracy	Method	k	Runtime (ms)	Accuracy	Method	k	Runtime (ms)	Accuracy
XYZ	16	0.32	0.21	XYZ	16	0.32	0.21	VV7	16	0.32	0.21
XYZ	32	0.61	0.44		32	0.61	0.44	AIL	32	0.61	0.44
Ours	9	0.47	0.26	Ours	9	0.47	0.26		9	0.47	0.26
Ours	18	0.99	0.77		18	0.99	0.77	Ours	18	0.99	0.77

You can compline rows if you want.
 \multipoord



You can combine rows if you want.
\multirow

- > Many prefer the third one.
- Personally speaking, the second one is straightforward enough in many cases (simpler is better).
- Choose what you like!

Method	k	Runtime (ms)	Accuracy	Method	k	Runtime (ms)	Accuracy	Method	k	$Runtime \ (ms)$	Accuracy
XYZ	16	0.32	0.21	XYZ	16	0.32	0.21	XV7	16	0.32	0.21
XYZ	32	0.61	0.44		32	0.61	0.44	AIL	32	0.61	0.44
Ours	9	0.47	0.26	Ours	9	0.47	0.26		9	0.47	0.26
Ours	18	0.99	0.77		18	0.99	0.77	Ours	18	0.99	0.77




## Multiple-level grouping

> You can group rows recursively.

Nearest Station	Store Name	Item	Price (JPY)
Hongo-Sanchome	Umite	Natsu-Ramen	700
		Maze Men X	850
	IBASA	Ramen	700
		Cold Ramen	650
Todai-mae	Yojinmen	Ramen	700
		Tsukemen	800

#### **Tables**

- Basics
- Row-oriented structure
- Row grouping
- Row hierarchization
- Column hierarchization
- Partial horizontal line (cmidrule)
- Column to row
- > Flowchart

## <u>Plots</u>

- Basics
- Bar chart to box plot
- Control parameter
- > Misc

- Row hierarchization = Deleting a row + Indentation
   In other words, horizontally-long + vertically-long
- > Row hierarchization may make a table easier to read.

Item	Type	Price $(\clubsuit)$	Floor
Pork	Meat	300	2
Beef	Meat	500	2
Tomato	Vegetables	100	3
$\operatorname{Corn}$	Vegetables	200	3
Cabbage	Vegetables	30	4

Item	Type	Price (¥)	Floo
Pork	Meat	300	2
Beef		500	2
Tomato	Vegetables	100	3
Corn		200	3
Cabbage		30	4

Item	Price (¥)	Floor
Meat		
Pork	300	2
Beef	500	2
Vegetables		
Tomato	100	3
Corn	200	3
Cabbage	30	4

#### A good table. Rows are interchangeable.

- Row hierarchization = Deleting a row + Indentation
   In other words, horizontally-long + vertically-long
- > Row hierarchization may make a table easier to read.

Item	Type	$\mathrm{Price}\;(\bigstar)$	Floor
Pork	Meat	300	2
Beef	Meat	500	2
Tomato	Vegetables	100	3
$\operatorname{Corn}$	Vegetables	200	3
Cabbage	Vegetables	30	4

Item	Type	Price $(\mathbf{¥})$	Floor
Pork	Meat	300	2
Beef		500	2
Tomato	Vegetables	100	3
Corn		200	3
Cabbage		30	4

Item	Price (¥)	Floor
Meat		
Pork	300	2
Beef	500	2
Vegetables		
Tomato	100	3
Corn	200	3
Cabbage	30	4



- Row hierarchization = Deleting a row + Indentation
   In other words, horizontally-long + vertically-long
- > Row hierarchization may make a table easier to read.

								Item	Price $(\mathbf{F})$	Floo
Item	Type	Price $(\textcircled{\textbf{H}})$	Floor	Item	Туре	Price $(\mathbf{¥})$	Floor	Meat		
Pork	Meat	300	2	Pork	Meat	300	2	Pork	300	2
Beef	Meat	500	2	Beef		000	2	Beef	500	2
Tomato	Vegetables	100	3	Tomato	Vegetables	100	3	Vegetables		
Corn	Vegetables	200	3	Corn		200	3	Tomato	100	3
Cabbage	Vegetables	30	4	Cabbage		30	4	Corn	200	3
								Cabbage	30	4
	Ir	ndenta	tion (	add whit	te snac	es in T	۵X)	<b>↔</b>		

Row hierarchization. In this case, this one is also good.



Indentation (add white spaces in TeX)

Row hierarchization. In this case, this one is also good.

- $\blacktriangleright$  Row hierarchization = Deleting a row + Indentation
  - ✓ In other words, horizontally-long  $\leftrightarrow$  vertically-long
- Row hierarchization may make a table easier to read.

				_						Item	Price $(\mathbf{F})$	Floor
Item	Type	Price $(\clubsuit)$	Floor	r	Item	Type	Price $(\clubsuit)$	Floor		Meat		
Pork	Meat	300	2		Pork	Meat	300	2		Pork	300	2
Beef	Meat	500	2		Beef		500	2		Beef	500	2
Tomato	Vegetables	100	3		Tomato	Vegetables	100	3		Vegetables		
Corn	Vegetables	200	3		Corn		200	3		Tomato	100	3
Cabbage	Vegetables	30	4		Cabbage		30	4		Corn	200	3
				-						Cabbage	30	4

Row-grouping and row-hierarchization are just for visualization
 Semantically, one should be able to convert the table to the original "stack of rows" from easily.

#### **Tables**

- Basics
- Row-oriented structure
- Row grouping
- Row hierarchization
- Column hierarchization
- Partial horizontal line (cmidrule)
- Column to row
- > Flowchart

## **Plots**

- Basics
- Bar chart to box plot
- Control parameter
- > Misc

#### **Column** hierarchization

> Columns cannot be grouped. But can be hierarchically summarized.



➢ Hierarchize "error".

No redundant information. Good!
 Again, this is for visualization. We can always back to the original form.

45

### **Column** hierarchization



- Hierarchize "error".
- No redundant information. Good!
   Again, this is for visualization. We can always back to the original form.

46

## Multiple column-hierarchization

> You can hierarchize columns multiple times.

With hierarchization, one needs to write a unit just once, making the table shorter

Temperature (°C		(°C)	Precipita	ation (mm)	Acces	s		
Prefecture	Max.	Avg.	Min.	August	December	By shinkansen	By plane	Density (persons/ $km^2$ )
Ishikawa	32	20	-1	179.8	304.7	$\checkmark$	$\checkmark$	267
Shizuoka	27	23	5	250.9	63.0	$\checkmark$	$\checkmark$	461
Okinawa	33	28	15	175.4	104.4		$\checkmark$	643

It's ok to mix usual columns and hierarchized columns

#### Multiple column-hierarchization

#### \begin{tabular}{@{}llllllll@{}} \toprule

& \multicolumn{3}{c}{Temperature (\$\mathrm{^\circ C}\$) } & \multicolumn{2}{c}{Precipitation (mm) } & \multicolumn{2}{c}{Access} & \\ \cmidrule(lr){2-4} \cmidrule(lr){5-6} \cmidrule(lr){7-8} Prefecture & Max. & Avg. & Min. & August & December & By shinkansen & By plane & Density (persons/\${km}^2\$) \\ \midrule

Ishikawa & 32 & 20 & -1 & 179.8 & 304.7 & \checkmark & \checkmark & 267 \\

Shizuoka & 27 & 23 & 5 & 250.9 & 63.0 & \checkmark & \checkmark & 461 \\

Okinawa & 33 & 28 & 15 & 175.4 & 104.4 & & \checkmark & 643 \\ \bottomrule

\end{tabular}

# a unit just once, making the table shorter

Tem		erature	$(^{\circ}C)$	Precipita	ation (mm)	Acces	s	
Prefecture	Max.	Avg.	Min.	August	December	By shinkansen	By plane	Density (persons/ $km^2$ )
Ishikawa	32	20	-1	179.8	304.7	$\checkmark$	$\checkmark$	267
Shizuoka	27	23	5	250.9	63.0	$\checkmark$	$\checkmark$	461
Okinawa	33	28	15	175.4	104.4		$\checkmark$	643

It's ok to mix usual columns and hierarchized columns

#### **Tables**

- Basics
- Row-oriented structure
- Row grouping
- Row hierarchization
- Column hierarchization
- Partial horizontal line (cmidrule)
- Column to row
- Flowchart

## **Plots**

- Basics
- Bar chart to box plot
- Control parameter
- > Misc

### Partial horizontal line (cmidrule)

- > Partial horizontal line (cmidrule) has two parameters.
  - ✓ Trims the ends of the line?: "1" (left) or "r" (right) or "1r" (both)
  - ✓ Range (e.g., "2-4")
- > By setting these params correctly, the table becomes much beautiful.



#### Partial horizontal line (cmidrule)

•	****	error				error				error				Error	
Method	Min.	Avg.	Max.	Method	Min.	Avg.	Max.	Method	Min.	Avg.	Max.	Method	Min.	Avg.	Max
Isomap	0.23	0.44	0.92	Isomap	0.23	0.44	0.92	Isomap	0.23	0.44	0.92	Isomap	0.23	0.44	0.92
LLE	0.10	0.73	1.82	LLE	0.10	0.73	1.82	LLE	0.10	0.73	1.82	LLE	0.10	0.73	1.82
	cmidrule(){2-4} cmidrule(1){2-4}			2-4}	cmidrule(r){2-4}				cmidrule(lr){2-4}						
	Trim left			Trim right			Trim left and right								
										0					0

Small differences... But god is in the details...

#### Partial horizontal line (cmidrule)

#### w/ proper params Temperature (°C) Precipitation (mm) Access Density (persons/ $km^2$ ) Avg. By shinkansen By plane Prefecture August December Max. Min. 179.8Ishikawa 3220-1 304.7 $\checkmark$ $\checkmark$ 267Shizuoka 2723 $\mathbf{5}$ 250.963.0 $\checkmark$ 461 $\checkmark$ Okinawa 332815175.4104.4643 $\checkmark$

w/o pr	oper	para	ms					
	Temp	erature	(°C)	Precipita	ation (mm)	Access	5	
Prefecture	Max.	Avg.	Min.	August	December	By shinkansen	By plane	Density (persons/ $km^2$ )
Ishikawa	32	20	-1	179.8	304.7	$\checkmark$	$\checkmark$	267
Shizuoka	27	23	5	250.9	63.0	$\checkmark$	$\checkmark$	461
Okinawa	33	28	15	175.4	104.4		$\checkmark$	643

#### **Tables**

- Basics
- Row-oriented structure
- Row grouping
- Row hierarchization
- Column hierarchization
- Partial horizontal line (cmidrule)
- Column to row
- Flowchart

### **Plots**

- Basics
- Bar chart to box plot
- Control parameter
- > Misc

This is not recommended, but moving a column to a row often makes the table clearer. Consider the following typical table.

Method	Dataset	Runtime (ms)
k-means	MNIST	10.2
	ImageNet	45.3
	Places	57.1
Ours	MNIST	8.3
	ImageNet	39.1
	Places	82.3

- > This table is good (rows are interchangeable), but...
  - ✓ If we compare methods in the same dataset, the values are far away (e.g., on MNIST, k-means (10.2) vs Ours (8.3)).
  - $\checkmark$  With more methods/datasets, the table becomes more vertical.

This is not recommended, but moving a column to a row often makes the table clearer. Consider the following typical table.

	Method	Dataset	Runtime (ms)	
	k-means	MNIST	10.2	
		ImageNet	45.3	
		Places	57.1	
	Ours	MNIST	8.3	
,		ImageNet	39.1	
		Places	82.3	

- > This table is good (rows are interchangeable), but...
  - ✓ If we compare methods in the same dataset, the values are far away (e.g., on MNIST, k-means (10.2) vs Ours (8.3)).
  - $\checkmark$  With more methods/datasets, the table becomes more vertical.

This is not recommended, but moving a column to a row often makes the table clearer. Consider the following typical table.

Dataset	Runtime (ms)
MNIST	10.2
ImageNet	45.3
Places	57.1
MNIST	8.3
ImageNet	39.1
Places	82.3
	Dataset MNIST ImageNet Places MNIST ImageNet Places

> This table is good (rows are interchangeable), but...

✓ If we compare methods in the same dataset, the values are far away (e.g., on MNIST, k-means (10.2) vs Ours (8.3)).

✓ With more methods/datasets, the table becomes more vertical.

		F	Restructure the probler Each row is (method Each row is (method	n. , dataset, , runtime	runtime) to for MNIST, runti	me for ImageNet,	runtime for Pla
Method	Dataset	Runtime (ms)	-				
k-means	MNIST	10.2			Runtime (ms) for	Runtime (ms) for	Runtime (ms) for
	ImageNet	45.3		Method	Dataset=MNIST	Dataset=ImageNet	Dataset=Places
	Places	57.1	Column-to-row	k-means	10.2	45.3	57.1
Ours	MNIST	8.3		Ours	8.3	39.1	82.3
	ImageNet	39.1					
	Places	82.3				Column	
			-		h	Runtime (ms)	
				Method	Dataset=MNIST	Dataset=ImageNet	Dataset=Places
				k-means	10.2	45.3	57.1
		Stil	l wordy? 🤗 👘	Ours	8.3	39.1	82.3





Seems clear. But there's no descriptions for "Dataset"
 This approach works only if it's acceptable not to show "Dataset"



Seems perfect? But there's no description for "Runtime".
 You need to write "This table reports runtime (ms)" in the caption.
 If a table consists of only one type of value, you may clarify the table by removing the type in the header and explaining the type in the caption.



Seems perfect? But there's no description for "Runtime".
 You need to write "This table reports runtime (ms)" in the caption.
 If a table consists of only one type of value, you may clarify the table by removing the type in the header and explaining the type in the caption.



The original cause is that there are two factors to focus (Method and Dataset).
If there are two factors to focus, the table would be "matrix-like".
In this case, plotting may be better (you can further improve this by box-plot!).

ImageNet

Places

MNIST

#### **Tables**

- Basics
- Row-oriented structure
- Row grouping
- Row hierarchization
- Column hierarchization
- Partial horizontal line (cmidrule)
- Column to row
- Flowchart

## **Plots**

- Basics
- Bar chart to box plot
- Control parameter
- > Misc

### Flowchart

- > When you have trouble creating tables, then...
- 1. Break up the table until it becomes the "good" table (e.g., satisfying stack-of-rows principal). It is OK if it is very vertically long.
- 2. Repeat the "row grouping" for the most important column. Do "row hierarchization" as needed.
- 3. Perform "column grouping" as necessary.
- 4. If it still does not fit, do "move a column to a row" and then do "row hierarchization". Delete redundant descriptions.
- 5. If it still does not fit well, consider using a plot.

#### **Tables**

- Basics
- Row-oriented structure
- Row grouping
- Row hierarchization
- Column hierarchization
- Partial horizontal line (cmidrule)
- Column to row
- > Flowchart

# <u>Plots</u>

#### Basics

- Bar chart to box plot
- Control parameter
- > Misc

#### Basics

> Three basic plots; line chart, box plot (bar chart), and scatter plot

- > Try these basic plots first. In many cases, that will suffice.
- Avoid using overly complex/sophisticated plots, as reviewers may have difficulty understanding them.



Line chart







Scatter plot

Bar chat is more natural, but a bar chart can be (and should be) converted to a box plot.



#### Basics

#### > Three basic plots; line chart, box plot (bar chart), and scatter plot



One control parameter (a continuous value, not discrete labels).
 Compare several functions (lines) of the parameter.



The font size should be large! Ideally, the same size as the main text.

#### **Tables**

- Basics
- Row-oriented structure
- Row grouping
- Row hierarchization
- Column hierarchization
- Partial horizontal line (cmidrule)
- Column to row
- Flowchart

### <u>Plots</u>

- Basics
- Bar chart to box plot
- Control parameter
- > Misc

#### Bar chart to box plot

- > A bar chart is the first choice, however,
  - ✓ A bar chart with error bars is basically better.
  - ✓ A box plot is further better.
- > Considering the following observations.

Accuracy of AlexNet [0.14 0.06 0.05 0.11 0.14 0.08 0.2 0.14 0.1 0.08 0.07 0.15 0.09 0.01 0.08 0.15 0.04 0.04 0.11 0.11]

Accuracy of VGG [ 0.06 0.08 0.07 0.06 0.07 0.03 0.04 0.02 0.03 0.06 0.01 0.04 0.04 0.05 0.04 0.03 0.04 0.09 0.09 0.05]

Accuracy of ResNet [ 0.32 0.06 0.16 0.22 0.45 0.13 0.25 0.51 0.27 0.32 0.06 0.43 0.31 0.28 0.18 0.16 0.22 0.35 0.46 0.27]

Accuracy of DenseNet [0.42 0.36 0.39 0.43 0.55 0. 0.49 0.12 0.08 0.35 0.01 0.58 0.39 0.46 -0.05 0.37 0.34 0.06 0.05 0.44]

Accuracy of EfficientNet [ 0.1 0.21 0.17 0.12 0.28 0.25 0.11 0.01 0.09 0.12 0.21 0.3 0.11 0.23 0.18 0.21 0.05 0.12 0.07 0.1 ]

#### Bar chart to box plot

- > A bar chart is the first choice, however,
  - ✓ A bar chart with error bars is basically better.
  - ✓ A box plot is further better.
- Considering the following observations.

Accuracy of AlexNet [0.14 0.06 0.05 0.11 0.14 0.9 0.2 0.14 0.1 0.02 0.15 0.00 0.01 0.00 0.15 0.04 0.01 0.01 Comparison over the discrete labels (methods).

Accuracy of VGG [ 0.06 0.08 0.07 0.06 0.07 0.03 0.04

Accuracy of ResNet

[0.32 0.06 0.16 0.22 0.45 0.13 0.25 0.51 0.27 0.32 0.06 0.43 0.31 0.28 0.18 0.16 0.22 0.35 0.46 0.27]

So we don't use a line chart.

Let's consider a bar chart first.

DI

Accuracy of DenseNet [0.42 0.36 0.39 0.43 0.55 0. 0.49 0.12 0.08 0.35 0.01 0.58 0.39 0.46 -0.05 0.37 0.34 0.06 0.05 0.44]

Accuracy of EfficientNet [ 0.1 0.21 0.17 0.12 0.28 0.25 0.11 0.01 0.09 0.12 0.21 0.3 0.11 0.23 0.18 0.21 0.05 0.12 0.07 0.1 ]

#### Bar chart to box plot

Comparison of the average values.
That's it. Not so much informative.

#### By a box plot, we can see more information with the same space!



With error bars, we can see the variation of values.
 More informative with the same space (area)
#### Bar chart to box plot

Comparison of the average values.That's it. Not so much informative.

#### By a box plot, we can see more information with the same space!



With error bars, we can see the variation of values.
More informative with the same space (area)

#### Bar chart to box plot

Comparison of the average values.
That's it. Not so much informative.

# By a box plot, we can see more information with the same space!



With error bars, we can see the variation of values.
 More informative with the same space (area)

# Box plot



# Box plot



M. Streit and N. Gehlenborg, "Bar charts and box plots", Nature Methods, 2014. https://www.nature.com/articles/nmeth.2807

# Violin plot?

Can be found in the seaborn package



- > A violin plot is considered as a more sophisticated visualization.
- > We can visualize the distribution itself.
- > I personally don't recommend a violin plot much.

# Violin plot: bad points

A violin plot automatically generates (interpolates) the distribution.
 It is dangerous, especially when #data is not enough.



> The reviewers may not understand the violin plot anyway.

# Violin plot: good points

➢ If the data is bi-modal or multi-modal, a violin plot is the only option.
Data: [10.1, 9.9, 10.1, 9.8, 10.2, 10.0, 20.1, 20.2, 20.0, 19.9, 20.1]



#### **Tables**

- Basics
- Row-oriented structure
- Row grouping
- Row hierarchization
- Column hierarchization
- Partial horizontal line (cmidrule)
- Column to row
- Flowchart

# <u>Plots</u>

- Basics
- Bar chart to box plot
- Control parameter
- > Misc

It often happens that we would like to show the two line-plots with the same control parameter.



It often happens that we would like to show the two line-plots with the same control parameter.
Inverse of runtime. Higher is better.







# Informative plot with a less space!Trade-off between the two curves.



# Informative plot with a less space! Trade-off between the two curves.



#### Precision-recall curve has the same structure



https://scikit-learn.org/stable/auto\_examples/model\_selection/plot\_precision\_recall.html

### **Tables**

- Basics
- Row-oriented structure
- Row grouping
- Row hierarchization
- Column hierarchization
- Partial horizontal line (cmidrule)
- Column to row
- Flowchart

# <u>Plots</u>

- Basics
- Bar chart to box plot
- Control parameter
- > Misc

# Put the grid on a logarithmic graph



88

# Legend

The order of the methods in the legend should match the lines as much as possible.





### Don't use dangerous charts



https://matplotlib.org/stable/gallery/pie\_and\_polar\_charts/pie\_features.html#sphx-glr-gallery-pie-and-polar-charts-pie-features-py https://en.wikipedia.org/wiki/Radar\_chart https://www.forbes.com/sites/naomirobbins/2012/06/07/trellis-plot-alternative-to-three-dimensional-bar-charts/?sh=7ed6126d7dab

# Continuous parameter or discrete labels

> Don't draw a line for data of discrete labels.



# What should we do to improve our paper?

- Do good research
  - ✓ 
     <sup></sup> Yes, of course! But we know it takes time.
- Improve a writing skill.
  - $\checkmark$   $\cong$  Yes, of course! But we know it takes time.
- Improve English (if you're a non-native English speaker).
   Yes, of course! But we know it takes time.
- Improve equations/tables/plots.
  - If the study of the study of

#### Schedule

Date (2024)	Contents	Presented by
Week 1, Apr 10	Introduction. Review of fundamental concepts	<del>Yusuke, Koya, Yuki, Jun</del>
Week 2, Apr 17	Equations and pseudo-codes	<del>Yusuke Matsui</del>
Week 3, Apr 24	Presentation	Koya Narumi
Week 4, May 1	Tables and plots	<del>Yusuke Matsui</del>
Week 5, May 8	Figures	Koya Narumi
Week 6, May 22	Videos	Koya Narumi
Week 7, May 29	Invited Talk 1	Dr. Yoshiaki Bando (AIST)
Week 8, June 5	Invited Talk 2	Prof. Katie Seaborn (Tokyo Tech)
Week 9, June 12	GitHub in depth	Yusuke Matsui
Week 10, June 19	Automation of research and research dissemination (Web, Cloud, Cl/CD)	Jun Kato
Week 11, June 26	Research community	Jun Kato
Week 12, July 3	3DCG illustrations	Yuki Koyama
Week 13, July 10	Final presentations	_