

有用なR分析技術

②最高の視覚化Rパッケージ “ggplot2” 1



ggplot2を使う理由は？

①とにかく絵がきれい！

→Rの基本関数を使うよりきれいな絵を作ることができる

②様々なオプションをいじることができる！

→軸の最大値、グラフの色、条件を付けて色を変更するなど様々なオプションの変更が可能
→全てのパラメータを暗記することは不可能なため、分からない部分が出てきたらグーグルして対応

③情報が豊富！

→代表的なRの視覚化パッケージのため、分からないことができた場合はグーグルすればすぐ情報を見つけることができる

④他のtidyverseパッケージとの連動も簡単！

→ggplot2などの強力な機能を持つパッケージと同じtidyverseパッケージのため連動が簡単

練習に使うデータの紹介

前回のファイルを使ってdplyr関数でAverageカラムを作ってください！

	Student Number	Name	Major	Midterm	Final
1	1301	Choi	Literature	50	45
2	1302	Park	Engineering	76	42
3	1303	Han	Literature	100	98
4	1304	Jin	Engineering	85	92
5	1305	Liu	Medicine	100	100
6	1306	Li	Engineering	86	100
7	1307	Yamaoka	Social Science	75	90
8	1308	Sirasaki	Natural Science	88	90
9	1309	Honda	Natural Science	92	88
10	1310	Yamada	Art	45	0
11	1311	Nguyen	Literature	100	92
12	1312	Lee	Medicine	85	77
13	1313	Hong	Engineering	82	90
14	1314	Hua	Natural Science	56	42
15	1315	Ma	Natural Science	87	78
16	1316	Okusora	Literature	75	62
17	1317	Tu	Engineering	83	55
18	1318	Satou	Art	90	82

```
> df_e1 <- df %>% dplyr::mutate(Average = (Midterm + Final)/2)
> df_e1
# A tibble: 18 × 6
  `Student Number` Name      Major      Midterm Final Average
      <dbl> <chr>      <chr>      <dbl> <dbl> <dbl>
1         1301 Choi      Literature      50     45    47.5
2         1302 Park      Engineering     76     42     59
3         1303 Han      Literature     100     98     99
4         1304 Jin      Engineering     85     92    88.5
5         1305 Liu      Medicine      100    100    100
6         1306 Li      Engineering     86    100     93
7         1307 Yamaoka  Social Science    75     90    82.5
8         1308 Sirasaki Natural Science    88     90     89
9         1309 Honda   Natural Science    92     88     90
10        1310 Yamada   Art              45      0    22.5
11        1311 Nguyen   Literature      100     92     96
12        1312 Lee      Medicine        85     77     81
13        1313 Hong     Engineering     82     90     86
14        1314 Hua      Natural Science    56     42     49
15        1315 Ma      Natural Science    87     78    82.5
16        1316 Okusora  Literature       75     62    68.5
17        1317 Tu      Engineering     83     55     69
18        1318 Satou    Art              90     82     86
```

基礎となるggplo2の文法

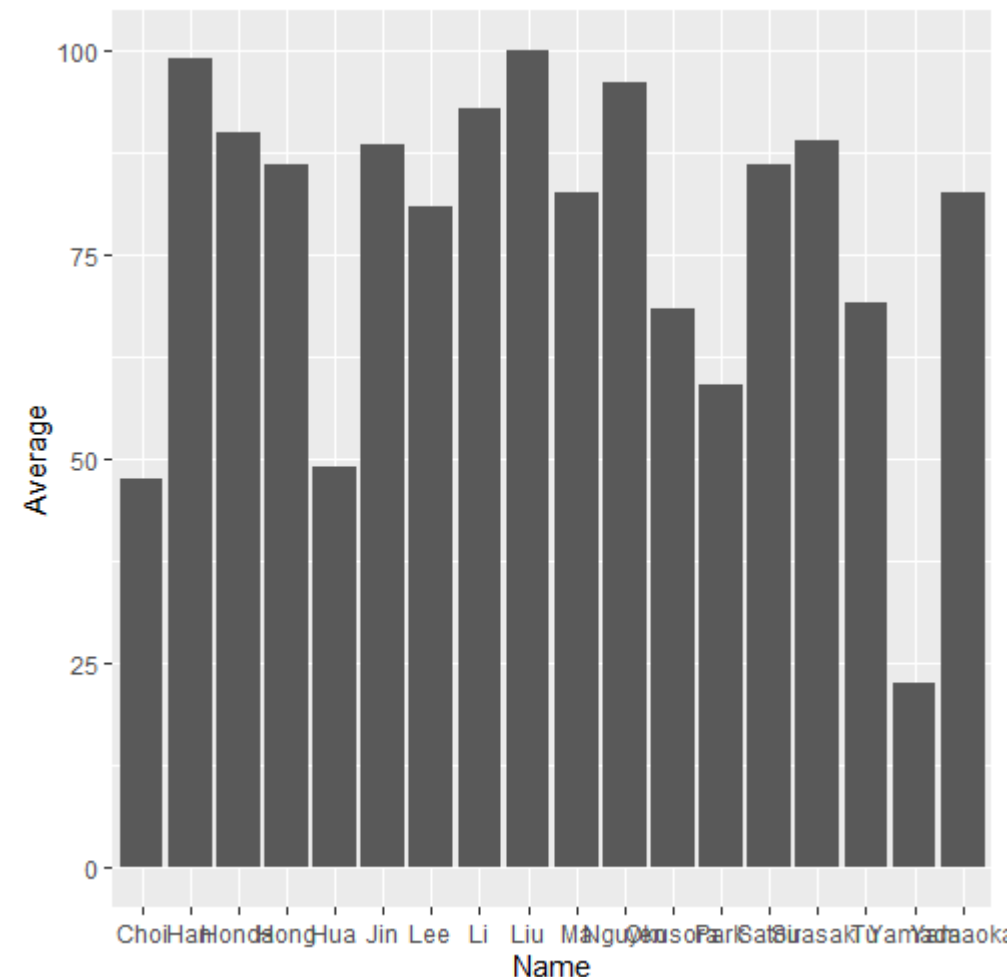
元データ

各軸を指定

```
> ggplot(df_e1, aes(x = Name, y = Average)) +  
+ geom_bar(stat = "identity")
```

Bar plotを作る

(stat = "identity")はデータを
修正せずそのまま使うという意味

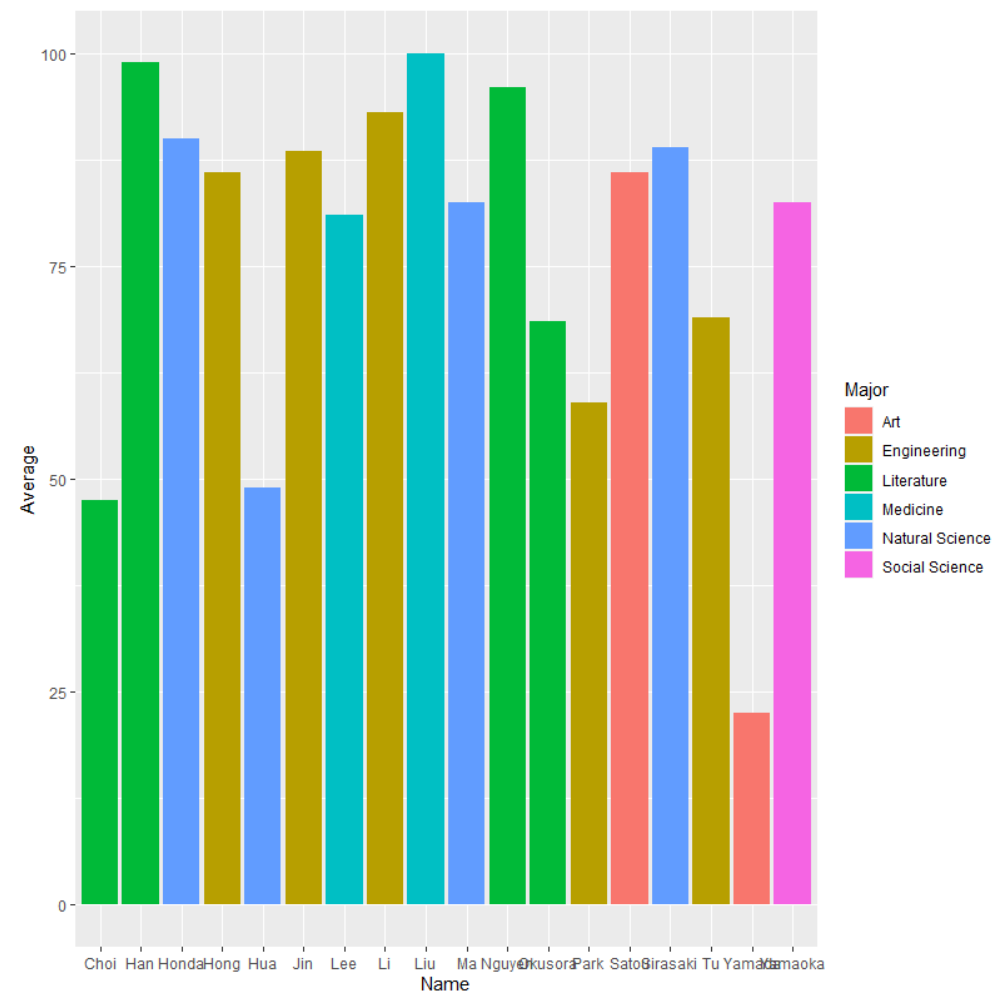


グラフセッティング

①カラーを付ける

```
> ggplot(df_e1, aes(x = Name, y = Average, fill = Major)) +  
+   geom_bar(stat = "identity")
```

Majorを基準に
して色を付ける



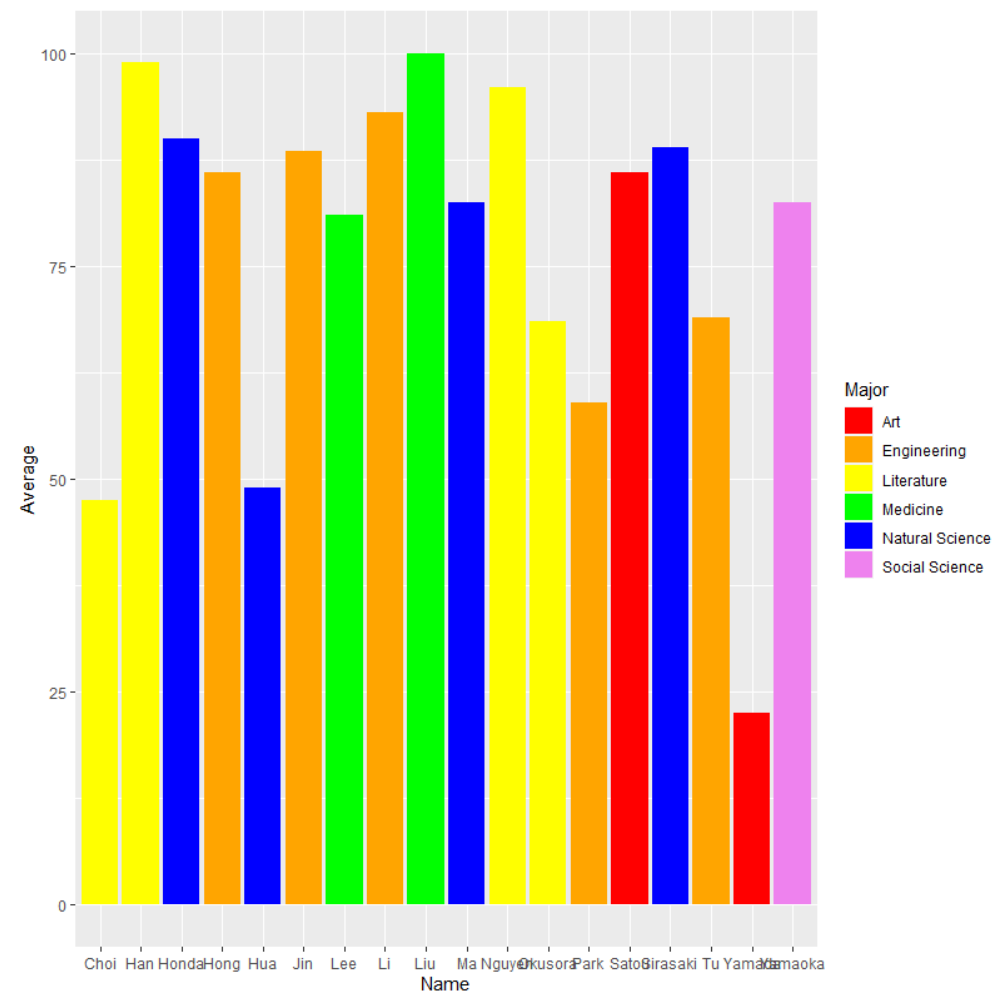
グラフセッティング

②凡例の順番と色を変更

凡例の順番を決める

```
> ggplot(df_e1, aes(x = Name, y = Average, fill = Major)) +  
+   geom_bar(stat = "identity") +  
+   scale_fill_discrete(breaks = c('Medicine', 'Engineering', 'Natural Science',  
+                                   'Social Science', 'Literature', 'Art')) +  
+   scale_fill_manual(values = c('red', 'orange', 'yellow',  
+                                   'green', 'blue', 'violet'))
```

任意の色名を入力

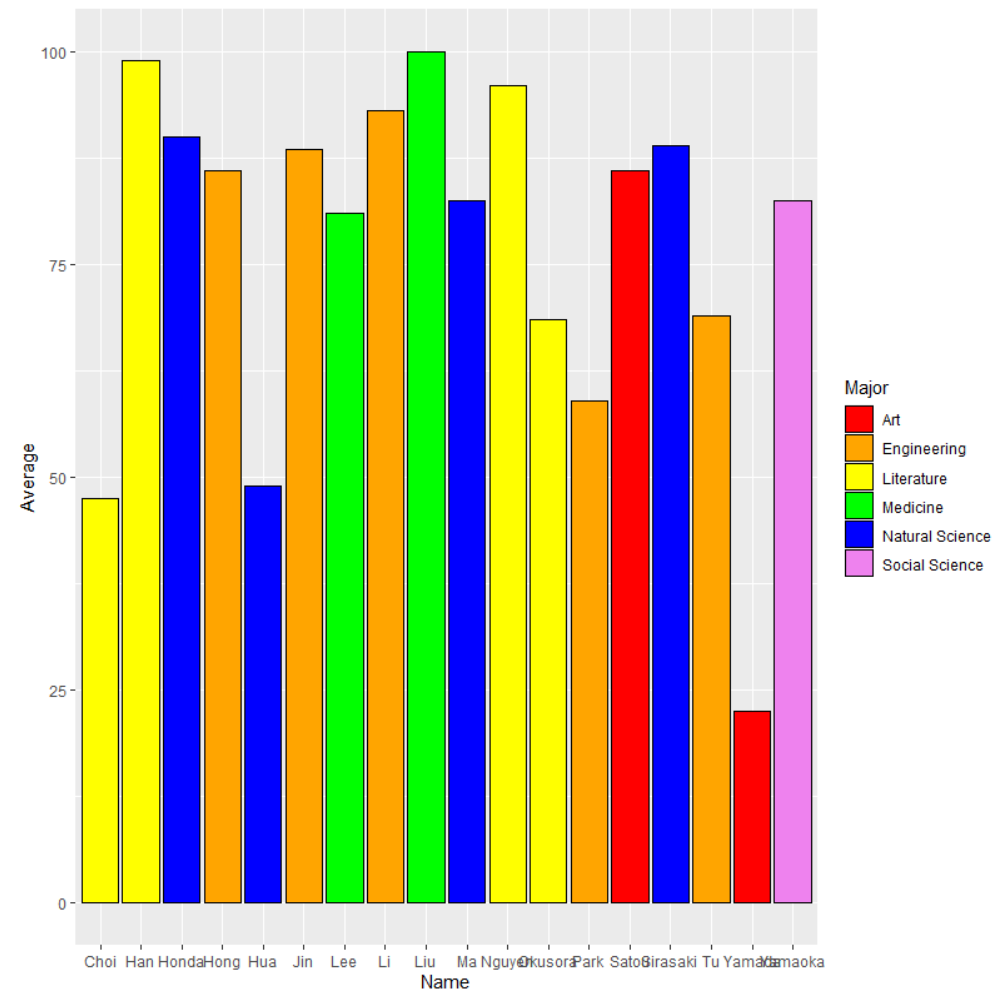


グラフセッティング

③棒の枠線を追加

各棒の枠線を決める

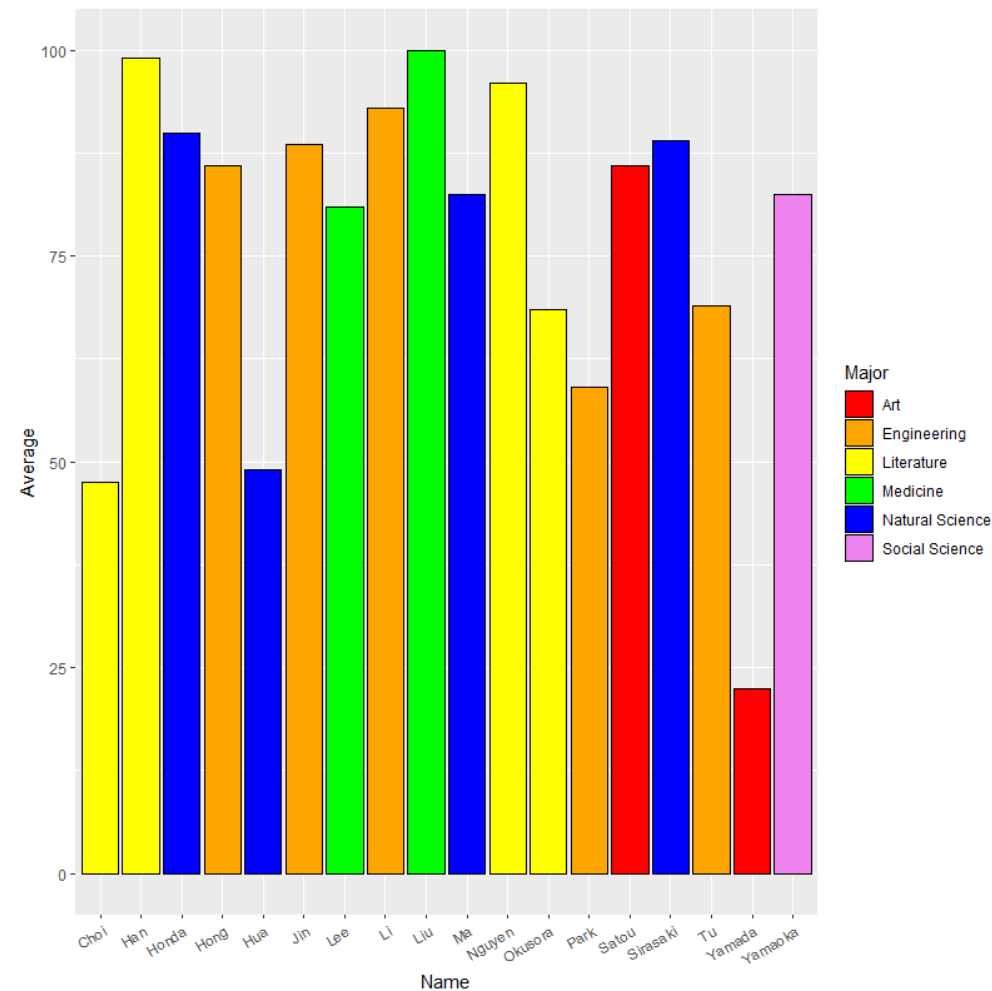
```
> ggplot(df_e1, aes(x = Name, y = Average, fill = Major)) +  
+   geom_bar(stat = "identity", color = 'black') +  
+   scale_fill_discrete(breaks = c('Medicine', 'Engineering', 'Natural Science',  
+   'Social Science', 'Literature', 'Art')) +  
+   scale_fill_manual(values = c('red', 'orange', 'yellow',  
+   'green', 'blue', 'violet'))
```



グラフセッティング

④ X軸の目盛を30度回す

```
> ggplot(df_e1, aes(x = Name, y = Average, fill = Major)) +  
+   geom_bar(stat = "identity", color = 'black') +  
+   scale_fill_discrete(breaks = c('Medicine', 'Engineering', 'Natural Science',  
+   'Social Science', 'Literature', 'Art')) +  
+   scale_fill_manual(values = c('red', 'orange', 'yellow',  
+   'green', 'blue', 'violet')) +  
+   scale_x_discrete(guide = guide_axis(angle = 30))
```



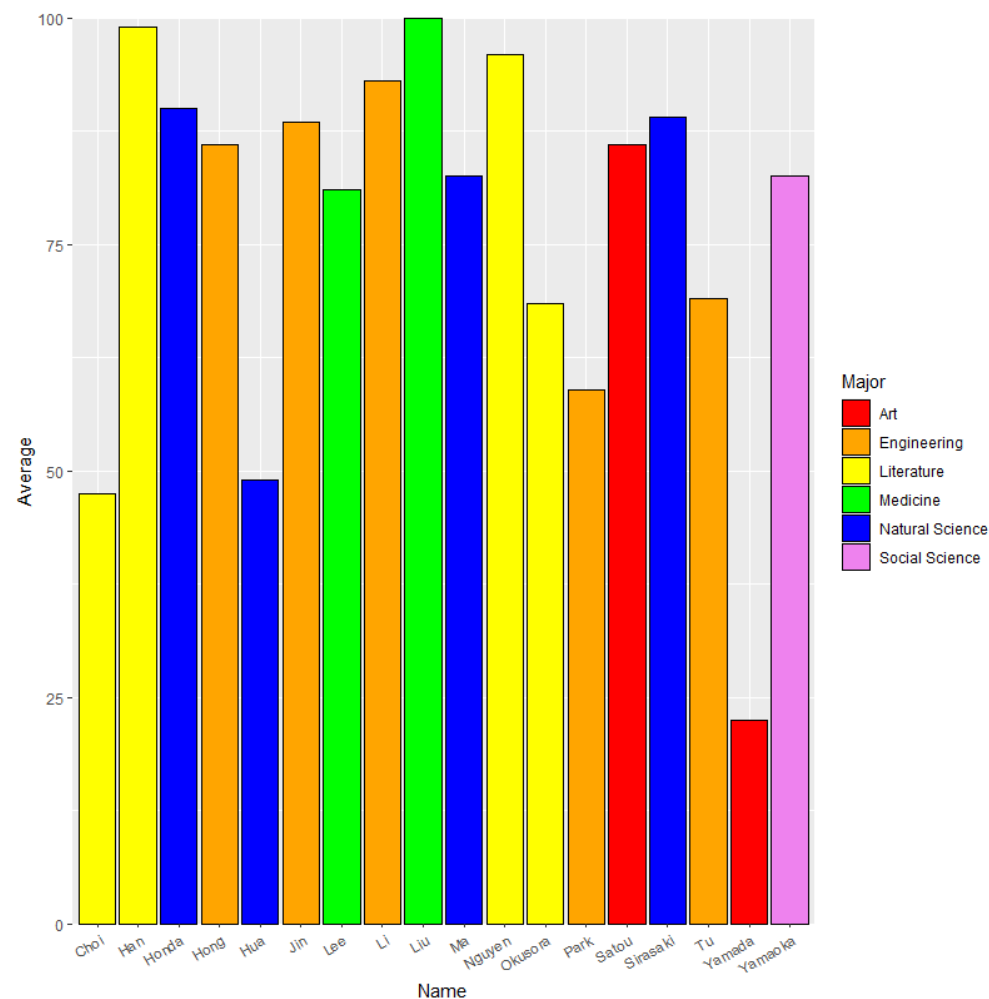
グラフセッティング

⑤ Y軸の範囲を設定・Y軸の最低値をX軸にくっつける

```
> ggplot(df_e1, aes(x = Name, y = Average, fill = Major)) +  
+   geom_bar(stat = "identity", color = 'black') +  
+   scale_fill_discrete(breaks = c('Medicine', 'Engineering', 'Natural Science',  
+   'Social Science', 'Literature', 'Art')) +  
+   scale_fill_manual(values = c('red', 'orange', 'yellow',  
+   'green', 'blue', 'violet')) +  
+   scale_x_discrete(guide = guide_axis(angle = 30)) +  
+   scale_y_continuous(limits = c(0, 100), expand = c(0, 0))
```

0から100までの範囲

空白を整理する

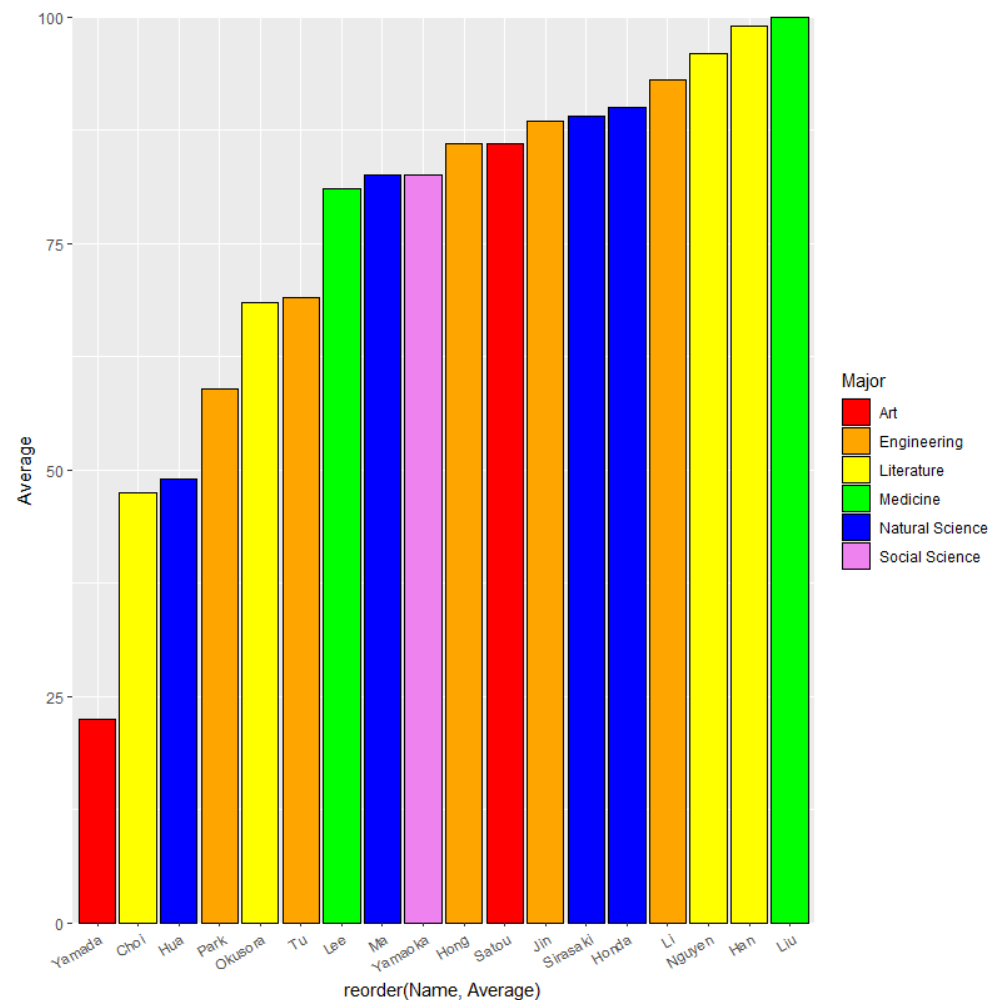


グラフセッティング

⑥各数値を整列する

NameをAverage基準に整列

```
> ggplot(df_e1, aes(x = reorder(Name, Average), y = Average, fill = Major)) +  
+   geom_bar(stat = "identity", color = 'black') +  
+   scale_fill_discrete(breaks = c('Medicine', 'Engineering', 'Natural Science',  
+   'Social Science', 'Literature', 'Art')) +  
+   scale_fill_manual(values = c('red', 'orange', 'yellow',  
+   'green', 'blue', 'violet')) +  
+   scale_x_discrete(guide = guide_axis(angle = 30)) +  
+   scale_y_continuous(limits = c(0,100), expand = c(0, 0))
```

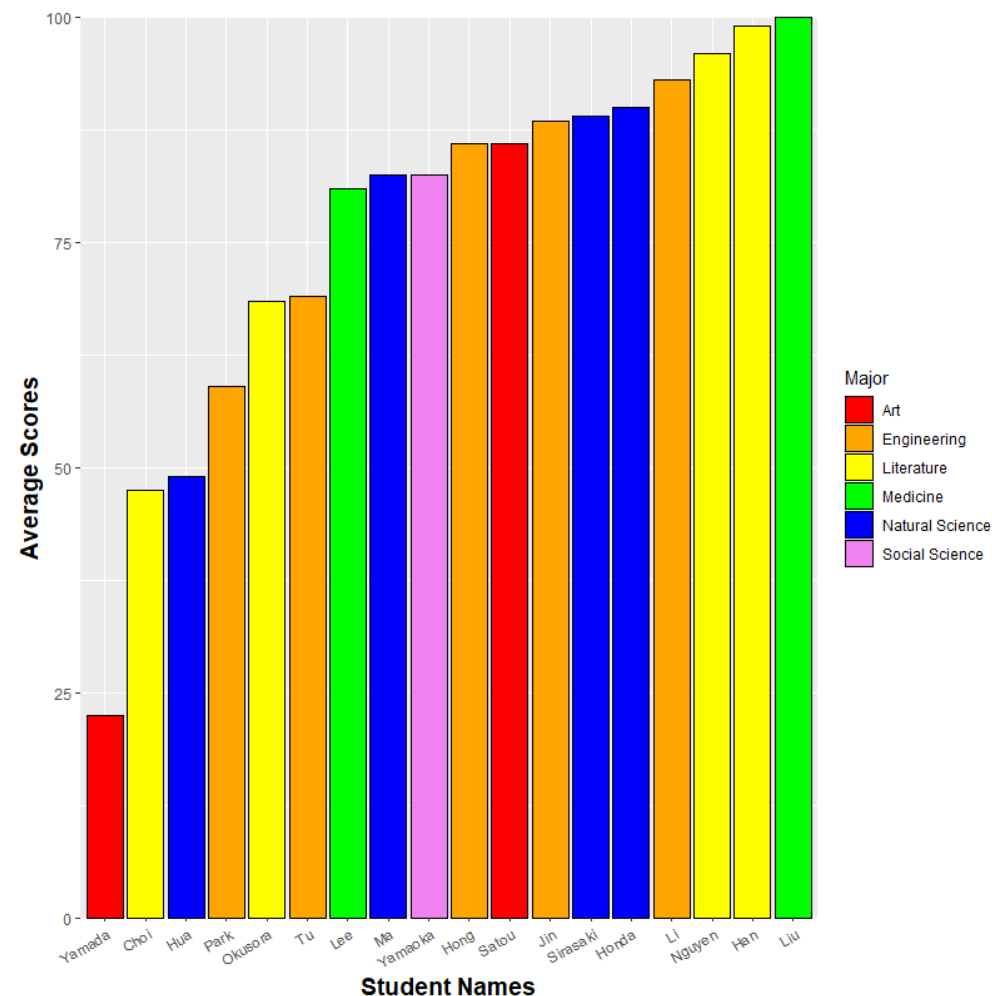


グラフセッティング

⑦軸名のデザインを変更する

```
> ggplot(df_e1, aes(x = reorder(Name, Average), y = Average, fill = Major)) +  
+   geom_bar(stat = "identity", color = 'black') +  
+   scale_fill_discrete(breaks = c('Medicine', 'Engineering', 'Natural Science',  
+   'Social Science', 'Literature', 'Art')) +  
+   scale_fill_manual(values = c('red', 'orange', 'yellow',  
+   'green', 'blue', 'violet')) +  
+   scale_x_discrete(guide = guide_axis(angle = 30)) +  
+   scale_y_continuous(limits = c(0,100), expand = c(0, 0)) +  
+   xlab('Student Names') +  
+   ylab('Average Scores') +  
+   theme(axis.title.x = element_text(size = 14, face = 'bold'),  
+   axis.title.y = element_text(size = 14, face = 'bold'))
```

軸名のサイズを14、太字に変更



グラフセッティング

⑧目盛のデザインを変更する

```
> ggplot(df_e1, aes(x = reorder(Name, Average), y = Average, fill = Major)) +  
+   geom_bar(stat = "identity", color = 'black') +  
+   scale_fill_discrete(breaks = c('Medicine', 'Engineering', 'Natural Science',  
+                                   'Social Science', 'Literature', 'Art')) +  
+   scale_fill_manual(values = c('red', 'orange', 'yellow',  
+                                   'green', 'blue', 'violet')) +  
+   scale_x_discrete(guide = guide_axis(angle = 30)) +  
+   scale_y_continuous(limits = c(0,100), expand = c(0, 0)) +  
+   xlab('Student Names') +  
+   ylab('Average Scores') +  
+   theme(axis.title.x = element_text(size = 14, face = 'bold'),  
+         axis.title.y = element_text(size = 14, face = 'bold'),  
+         axis.text.x = element_text(size = 8, face = 'bold'),  
+         axis.text.y = element_text(size = 12, face = 'bold'))
```

目盛のサイズを8もしくは12、
太字に変更

