# Seaborn မှ Heat Map အကြောင်း မိတ်ဆက်

Seaborn မှာ ပါတဲ့ Heat Map အကြောင်း မိတ်ဆက်ပေးချင်ပါတယ်။ Code တစ်ကြောင်းစီကို ရှင်းပြထားပြီး Heat Map သဘောတရားတွေကို ရှင်းပြထားပါတယ်။

ကြိုပြီးတော့ အကြမ်းဖျဉ်း ရှင်းပြရင် ကိန်းဂဏန်းတွေကို လိုက်ဖတ်ပြီး ဘယ်နေရာမှာ အရေအတွက် များတယ်။ ဘယ်နေရာမှာ အရေအတွက် နည်းတယ်ဆိုတာကို လိုက်စဉ်းစားနေမဲ့အစား အရောင်တွေကို ကြည့်ရုံနဲ့ အလွယ်တကူ သိနိုင်အောင်လုပ်ထားတာပါ။
ဥပမာ အရောင် အရင့်ဆုံးက တန်ဖိုး အများဆုံးဆိုရင် အရောင်ရင့်ရင်တွေ ဘယ်နေရာမှာ ရှိသလဲဆိုတာကြည့်ရုံနဲ့ ဘယ်နေရာမှာ တန်ဖိုး အများဆုံးတွေ ရှိတယ်ဆိုတာကို အလွယ်တကူ သိနိုင်ပါတယ်။

numpy ၊ matplotlib.pyplot နဲ့ seaborn ကို import လုပ်ပါတယ်။

In [1]:
```
%matplotlib inline

import numpy as np; np.random.seed(0)
import seaborn as sns; sns.set()
import matplotlib.pyplot as plt
```

## Data Loading

ဒေတာတွေကို တင်ပါတယ်။ seaborn ထဲမှာဒေတာတွေ ပါ ပြီးသား ပါပြီးသား မို့လို့ တခြား ဒေတာဖိုင်တွေကို ရှာဖွေနေစရာမလို့တော့ပါဘူး။ ဒေတာတွေ တက်လာဘို့ (loading ဖြစ်ဘို့) internet connection လိုပါတယ်။

`flights = sns.load_dataset("flights")` က ဒေတာတွေကို load လုပ်တာပါ။ ဒေတာတွေက ဘယ်လိုပုံစံနဲ့ ရှိသလဲဆိုတာကို သိဘို့ ပထမဆုံး(၅)ကြောင်းကို `flights.head()` နဲ့ ကြည့်ပါတယ်။ ဒေတာတွေက ဒေါင်လိုက်ရှိနေတာပါ။ ဒေတာတွေက ဒေါင်လိုက်ရှိနေတာပါ။ 1949 နောက်ပိုင်း ခုနှစ်တွေမှာ လေယာဉ်ပျံစီးသူ အရေအတွက် လအလိုက် သိမ်းဆည်းထားတဲ့ ဒေတာပါ။

In [2]:
```
flights = sns.load_dataset("flights")
flights.head()
```

Out[2]:

| | year | month | passengers |
|---|---|---|---|
| **0** | 1949 | January | 112 |
| **1** | 1949 | February | 118 |
| **2** | 1949 | March | 132 |
| **3** | 1949 | April | 129 |
| **4** | 1949 | May | 121 |

## Pivot Table

ခုနှစ်တွေ ဒေါင်လိုက်ရှိနေတာကို `flights.pivot("month", "year", "passengers")` သုံးပြီးတော့ အောက်မှာပြထားသလို ဇယားပုံစံဖြစ်အောင်လုပ်ပါတယ်။

In [3]:
```
flights = flights.pivot("month", "year", "passengers")
flights
```

Out[3]:

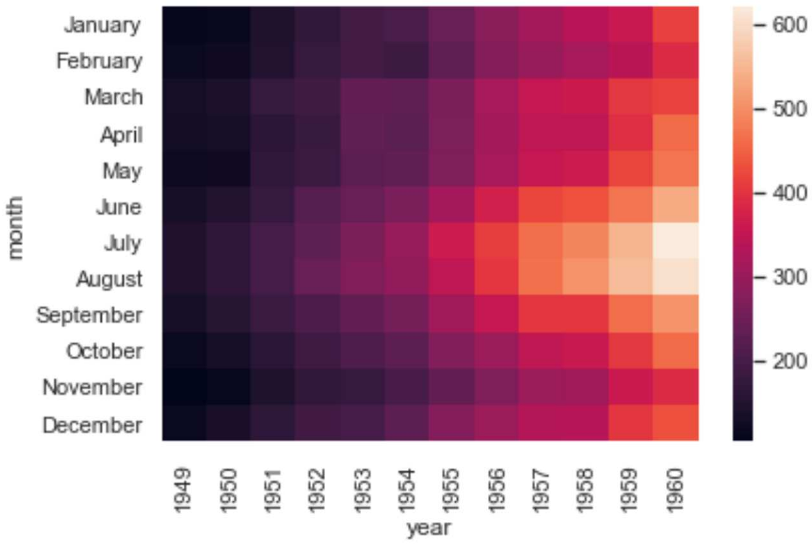
| year<br>month | 1949 | 1950 | 1951 | 1952 | 1953 | 1954 | 1955 | 1956 | 1957 | 1958 | 1959 | 1960 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **January** | 112 | 115 | 145 | 171 | 196 | 204 | 242 | 284 | 315 | 340 | 360 | 417 |
| **February** | 118 | 126 | 150 | 180 | 196 | 188 | 233 | 277 | 301 | 318 | 342 | 391 |
| **March** | 132 | 141 | 178 | 193 | 236 | 235 | 267 | 317 | 356 | 362 | 406 | 419 |
| **April** | 129 | 135 | 163 | 181 | 235 | 227 | 269 | 313 | 348 | 348 | 396 | 461 |
| **May** | 121 | 125 | 172 | 183 | 229 | 234 | 270 | 318 | 355 | 363 | 420 | 472 |
| **June** | 135 | 149 | 178 | 218 | 243 | 264 | 315 | 374 | 422 | 435 | 472 | 535 |
| **July** | 148 | 170 | 199 | 230 | 264 | 302 | 364 | 413 | 465 | 491 | 548 | 622 |
| **August** | 148 | 170 | 199 | 242 | 272 | 293 | 347 | 405 | 467 | 505 | 559 | 606 |
| **September** | 136 | 158 | 184 | 209 | 237 | 259 | 312 | 355 | 404 | 404 | 463 | 508 |
| **October** | 119 | 133 | 162 | 191 | 211 | 229 | 274 | 306 | 347 | 359 | 407 | 461 |
| **November** | 104 | 114 | 146 | 172 | 180 | 203 | 237 | 271 | 305 | 310 | 362 | 390 |
| **December** | 118 | 140 | 166 | 194 | 201 | 229 | 278 | 306 | 336 | 337 | 405 | 432 |

# Heat Map

`sns.heatmap(flights)` နဲ့ heat map ဂရပ်ဆွဲပါတယ်။

In [4]:
```
sns.heatmap(flights)
```

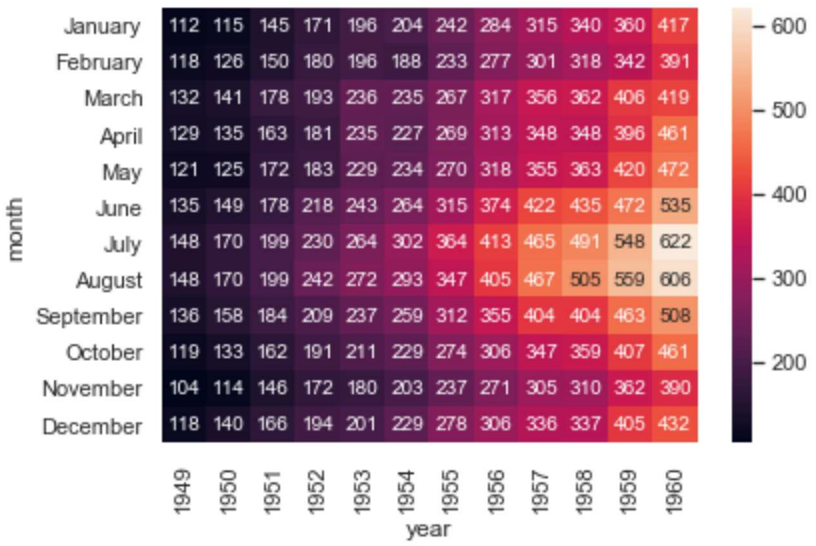
Out[4]: `<matplotlib.axes._subplots.AxesSubplot at 0xba25f28>`




ပေါ်လာတဲ့ heat map ဂရပ်က ဘာမှလည်း ဖတ်လို့ မရပါဘူး။ အဓိပ္ပာယ်လည်း ဖော်လို့ မရပါဘူး။ အဲဒါကြောင့် ဖတ်လို့ရအောင် အနည်းငယ် ပြင်ရပါမယ်။ အကွက်ထဲမှာ အရောင်တွေနဲ့ အတူ လေယာဉ်ပျံစီးသူ ခရီးသည် အရေအတွက်ကို တွဲပြီးတော့ ဖော်ပြအောင် `annot=True` ကို သုံးပါတယ်။

In [5]:
```
sns.heatmap(flights, annot=True, fmt="d")
```

Out[5]: <matplotlib.axes._subplots.AxesSubplot at 0xc18e780>

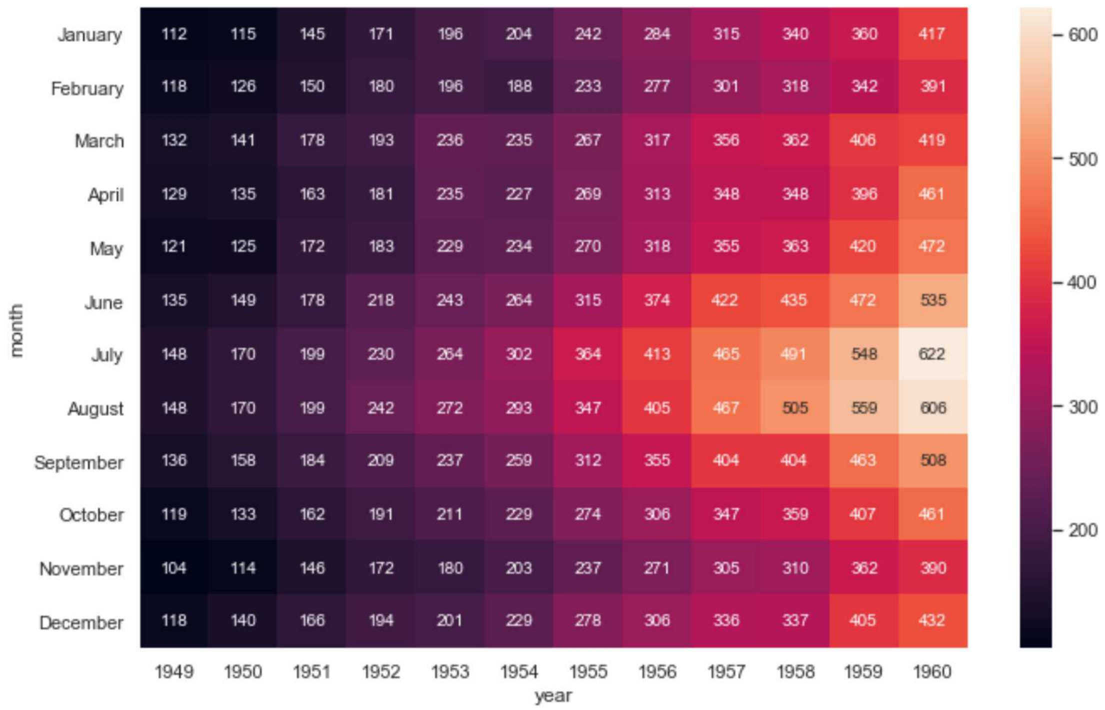



# ဂရပ် အရွယ်အစား(figsize)

heat map ဂရပ် အရွယ်အစားက သေးနေလို့ လိုချင်တဲ့ အရွယ်အစားရအောင် `plt.figure(figsize = (11,7))` ကို သုံးပြီးတော့ 11 x 7 ဖြစ်အောင် လုပ်ပါတယ်။

In [6]:
```
plt.figure(figsize = (11,7))
sns.heatmap(flights, annot=True, fmt="d")
```

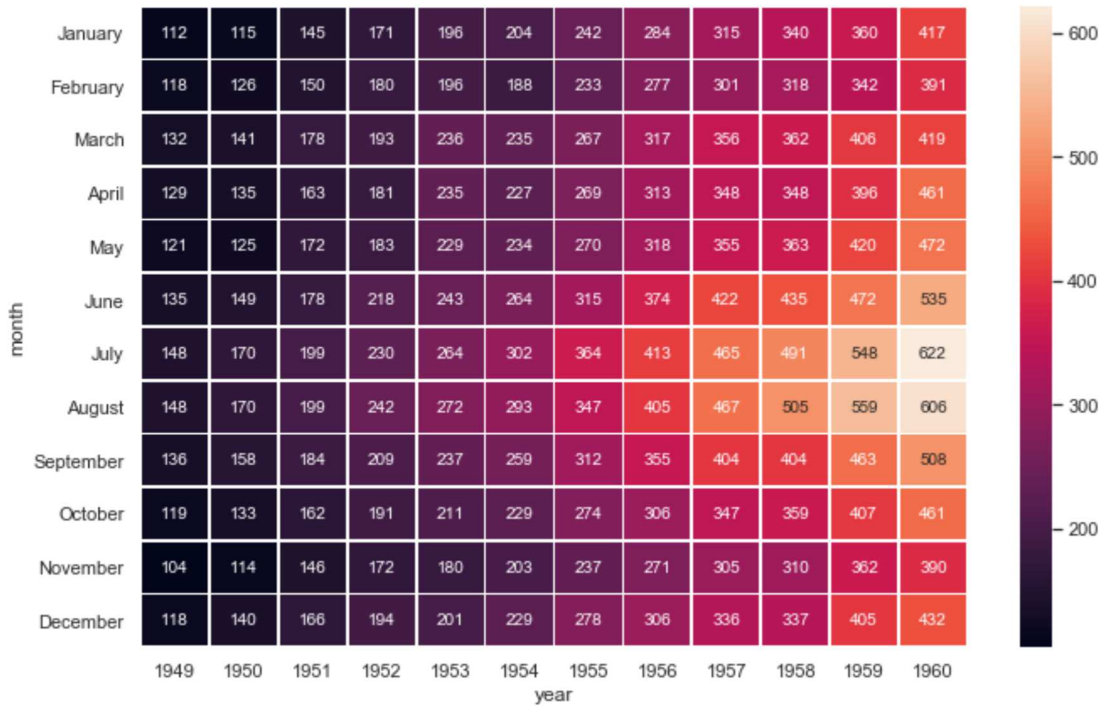
Out[6]: <matplotlib.axes._subplots.AxesSubplot at 0xc747208>



ကြည့်လို့ မကောင်းသေးတာကြောင့် အကွက်ကလေးတွေ တစ်ခုနှင့် တစ်ခုအကြားမှာ အဖြူရောင်လိုင်းကလေးတွေ ပေါ်လို့ `linewidth= 0.75` ကို သုံးပါတယ်။

```
In [7]: plt.figure(figsize = (11,7))
        sns.heatmap(flights, annot=True, fmt="d", linewidth= 0.75)
```

```
Out[7]: <matplotlib.axes._subplots.AxesSubplot at 0xcab3940>
```




အခုဆို heat map ကို ဖတ်လို့ ရပါပြီ။ အနည်းဆုံး ခရီးသည်အရေအတွက် ရှိတဲ့လ က 104 ဖြစ်ပြီးတော့ အများဆုံး ခရီးသည်အရေအတွက် ရှိတဲ့လက 622 ဖြစ်ပါတယ်။ ခရီးသည်အရေအတွက် အနည်းဆုံးလကို အနက်ရောင် ( အရောင်အရင့်ဆုံး) နဲ့ ဖော်ပြပြီးတော့ ခရီးသည်အရေအတွက် အများဆုံးလကို အရောင်အဖျော့ဆုံး နဲ့ ဖော်ပြပါတယ်။ ဘေးမှာ ကာလစကေးကိုလည်း ဖော်ပြထားပါတယ်။

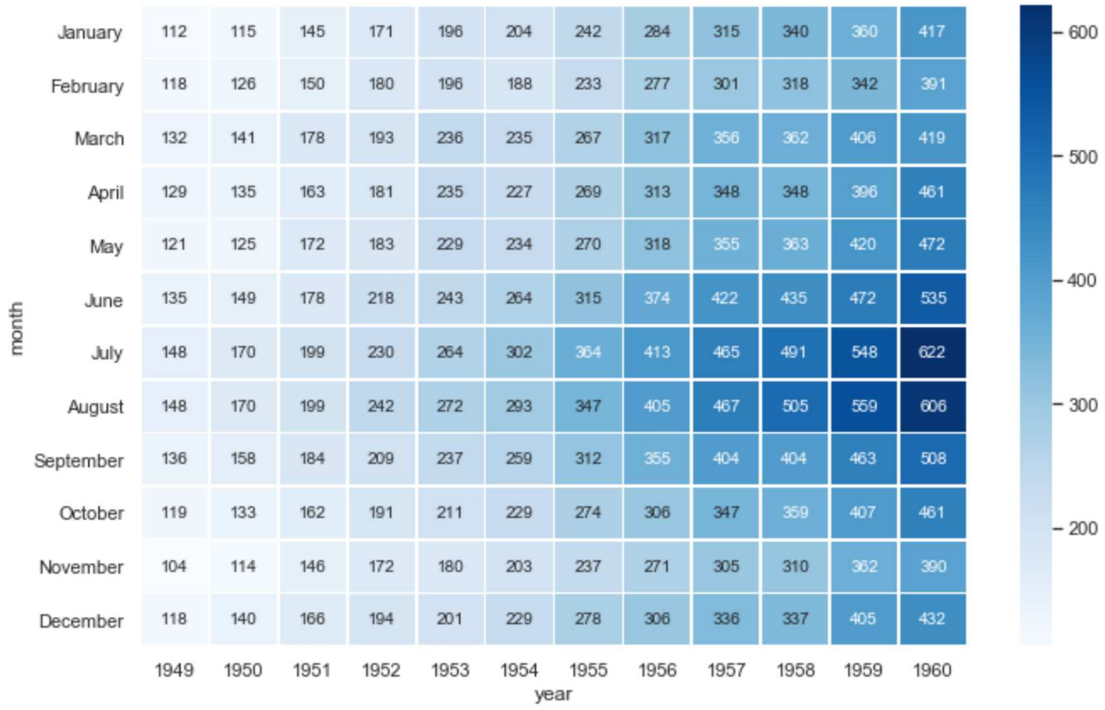
ကာလစကေးက အရောင်တွေကို စိတ်ထဲမှာ မှတ်ပြီးတော့ ဘယ်လတွေမှာ ခရီးသည်အရေအတွက် နည်းတယ်။ များတယ်ဆိုတာကို တစ်ချက်ကြည့်ရုံနဲ့ သိသာနိုင်ပါတယ်။ ဖိုတိုဂရပ်ဖစ် မန်မိုရီ ရှိတဲ့သူတွေဆို ဘယ်နှစ်ဘယ်လမှာ ခရီးသည်များတယ် နည်းတယ်ဆိုတာကို အလွတ်နီးပါး ပြောနိုင်လောက်တယ်။

## အရောင် ပြောင်းခြင်း

အနီရောင်နဲ့ အနက်ရောင် မကြိုက် တခြား အရောင်တွေ ပြောင်းလို့ ရပါတယ်။

In [8]:
```python
plt.figure(figsize = (11,7))
sns.heatmap(flights, annot=True, fmt="d", linewidth= 0.75,
            cmap = 'Blues')
```

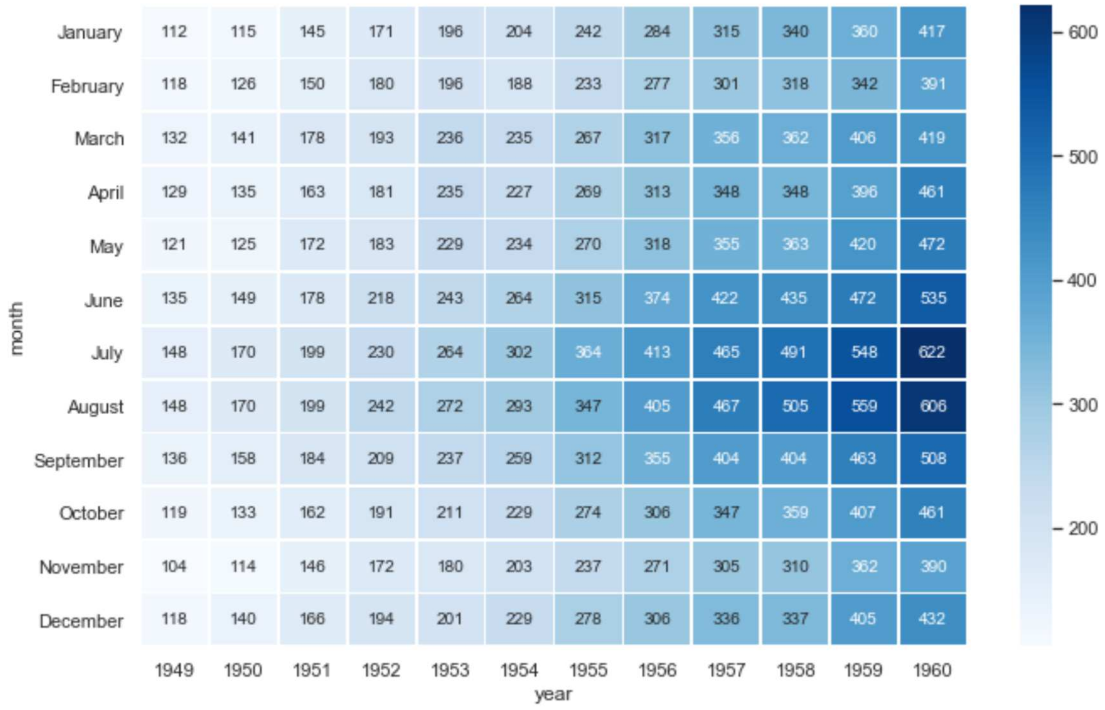
Out[8]: <matplotlib.axes._subplots.AxesSubplot at 0xc751710>




အပြာရောင် ပြောင်းလိုက်တဲ့အခါ သတိပြုဘို့ အချက် အများဆုံး ခရီးသည်အရေအတွက်များတဲ့ အကွက်တွေက အရောင် အရင်ဆုံးပါ။ အမြဲတမ်း တန်ဖိုးအများက အရောင် အရင့်ဆုံး ဖြစ်ရမယ်ရယ်လို့ မှတ်လို့ မရပါဘူး။

In [9]:
```
plt.figure(figsize = (11,7))
sns.heatmap(flights, annot=True, fmt="d",
            linewidth= 0.75, cmap = 'Blues', )
```

Out[9]: <matplotlib.axes._subplots.AxesSubplot at 0xcfc4c18>




# ကာလာစကေး ချိန်ခြင်း

ကာလာစကေးကို မိမိကြည့်ချင်တဲ့ တန်ဖိုး သို့မဟုတ် အရေအတွက် ထဲပဲထားလို့ ရပါတယ်။ `vmin = 0` , `vmax = 200` ကို သုံးပြီးတော့ ပြောင်းရပါတယ်။ vmin က အနည်းဆုံးပါ။ vmax က အများဆုံးပါ။
ကာလာစကေးက 0 နှင့် 200အကြားမို့ 0 ကနေ 200 အထိသာ အရောင်ပြောင်းပါတယ်။ 200 ကျော်တာနဲ့ တစ်ပြိုင်နက် တစ်ရောင်တည်းဖြစ်သွားပါပြီ။

```
In [10]: plt.figure(figsize = (11,7))
         sns.heatmap(flights, annot=True, fmt="d", linewidth= 0.75,
                     vmin = 0, vmax = 200)
```

Out[10]: <matplotlib.axes._subplots.AxesSubplot at 0xd7b66d8>

အောက်ကပုံရဲ့ ကာလာစကေးက 200 နှင့် 400 အကြားပါ။ အောက်ကပုံရဲ့ ကာလာစကေးက 200 နှင့် 400 အကြားပါ။ 200 နှင့် 200 အောက် အရေအတွက် ရှိတဲ့အကွက်တွေအားလုံးက အမဲရောင်တစ်ရောင်ထဲမှာ ဘာလို့လဲဆိုရင် ကာလာစကေးက 200 က စလို့ပါ။ ထို့အတူ 400 နှင့် 400 အထက်က အရောင်တွေအားလုံးက တူညီကြပါတယ်။ ဘာလို့လဲဆိုရင် ကာလာစကေးရဲ့ အမြင့်ဆုံးက 400 ဖြစ်လို့ပါ။ 400 ထက်ကျော်တာနဲ့ ကာလာစကေးရဲ့ အမြင့်ဆုံးအရောင်ကိုပဲ ဖော်ပြတော့ပါတယ်။

In [11]:
```
plt.figure(figsize = (11,7))
sns.heatmap(flights, annot=True, fmt="d", linewidth= 0.75,
            vmin = 200, vmax = 400)
```

Out[11]: <matplotlib.axes._subplots.AxesSubplot at 0xd532b70>

vmin နှင့် vmax ကို ထည့်မပေးထားရင် ဒေတာတွေအားလုံးခြုံမိအောင် ကာလာစကေးက အမြင့်ဆုံးက နေ အနိမ့်ဆုံးတန်ဖိုးထိအောင် ဖော်ပြပါလိမ့်မယ်။

In [ ]: