

导入pandas、warnings模块并控制小数位数

```
In [1]: import pandas as pd
import warnings
warnings.filterwarnings("ignore")
pd.options.display.float_format='{:.2f}'.format
```

读取记账数据

```
In [2]: df_income = pd.read_excel("4170421-2023年1~6月收入记账凭证.xlsx")
df_income.fillna(0,inplace=True) #对空值进行填充
df_income
```

```
Out[2]:
```

	科目代码	科目名称	客户名称	日期	凭证号	摘要	借方	贷方
0	5401	开票收入	腾讯公司	44933	1	销售收入	10000	20000
1	5402	无票收入	阿里巴巴集团	44940	2	服务费	3000	5000
2	5401	开票收入	京东集团	44947	3	销售收入	15000	25000
3	5403	无票收入	美团点评	44954	4	租赁收入	4000	6000
4	5402	无票收入	百度公司	44961	5	服务费	3500	5500
5	5401	开票收入	滴滴出行	44968	6	销售收入	18000	28000
6	5403	无票收入	快手科技	44975	7	租赁收入	4500	6500
7	5402	无票收入	小米公司	44982	8	服务费	4000	6000
8	5401	开票收入	华为集团	44989	9	销售收入	21000	31000
9	5403	无票收入	爱奇艺公司	44996	10	租赁收入	5500	7500

读取发票明细数据

```
In [3]: df_invo = pd.read_excel("4170421-2023年1~6月发票明细.xlsx")
df_invo.fillna(0,inplace=True) #对空值进行填充
df_invo
```

Out[3]:

	发票代码	发票号码	购方名称	商品名称	开票日期	购方税号	金额	税率	税额	作废
0	23456	1	腾讯公司	手机	44933	111111000000000000	2000	0.16	320.00	否
1	123456	2	阿里巴巴集团	电脑	44940	222222222222222016	5000	0.16	800.00	否
2	123456	3	京东集团	空调	44947	333333333333332992	2500	0.16	400.00	否
3	123456	4	美团点评	房屋租赁费	44954	444444444444444032	6000	0.16	960.00	否
4	123456	5	百度公司	服务费	44961	555555555555555008	5500	0.16	880.00	否
5	123456	6	滴滴出行	销售收入	44968	666666666666665984	28000	0.16	4480.00	否
6	123456	7	快手科技	租赁收入	44975	777777777777777024	6500	0.16	1040.00	否
7	123456	8	小米公司	服务费	44982	888888888888888064	6000	0.16	960.00	否
8	123456	9	华为集团	销售收入	44989	999999999999998976	31000	0.16	5760.00	否
9	123456	10	爱奇艺公司	租赁收入	44996	AAAAABBBBBCCCCCCC	7500	0.16	1289.00	是

筛选开票收入，计算凭证金额

```
In [4]: df_income_computer = df_income[df_income['科目名称'].str.contains('开票收入')]
df_income_computer['凭证金额']=df_income_computer['贷方']-df_income_computer['借方']
df_income_computer
```

Out[4]:

	科目代码	科目名称	客户名称	日期	凭证号	摘要	借方	贷方	凭证金额
0	5401	开票收入	腾讯公司	44933	1	销售收入	10000	20000	10000
2	5401	开票收入	京东集团	44947	3	销售收入	15000	25000	10000
5	5401	开票收入	滴滴出行	44968	6	销售收入	18000	28000	10000
8	5401	开票收入	华为集团	44989	9	销售收入	21000	31000	10000

筛选发票数据

对开票系统的发票明细数据进行筛选，把作废的数据删除，只保留有效发票

```
In [5]: df_invo_choose = df_invo[df_invo['作废']=='否']
df_invo_choose
```

Out[5]:

	发票代码	发票号码	购方名称	商品名称	开票日期	购方税号	金额	税率	税额	作废
0	23456	1	腾讯公司	手机	44933	111111000000000000	2000	0.16	320.00	否
1	123456	2	阿里巴巴集团	电脑	44940	222222222222222016	5000	0.16	800.00	否
2	123456	3	京东集团	空调	44947	333333333333332992	2500	0.16	400.00	否
3	123456	4	美团点评	房屋租赁费	44954	444444444444444032	6000	0.16	960.00	否
4	123456	5	百度公司	服务费	44961	555555555555555008	5500	0.16	880.00	否
5	123456	6	滴滴出行	销售收入	44968	666666666666665984	28000	0.16	4480.00	否
6	123456	7	快手科技	租赁收入	44975	777777777777777024	6500	0.16	1040.00	否
7	123456	8	小米公司	服务费	44982	888888888888888064	6000	0.16	960.00	否
8	123456	9	华为集团	销售收入	44989	999999999999998976	31000	0.16	5760.00	否

按客户名称汇总凭证金额

从客户维度比骄傲收入记账凭证和发票明细数据

```
In [6]: #按客户名称汇总凭证金额
dfincome = pd.pivot_table(df_income_computer, index=['客户名称'], values=['凭证金额'],
dfincome
```

Out[6]: **客户名称 凭证金额**

0	京东集团	10000
1	华为集团	10000
2	滴滴出行	10000
3	腾讯公司	10000

```
In [7]: # 按购方名称汇总开票金额
dfinvo = pd.pivot_table(df_invo_choose,index=['购方名称'],values=['金额'],aggfunc='sum')
dfinvo
```

Out[7]: **购方名称 金额**

0	京东集团	2500
1	华为集团	31000
2	小米公司	6000
3	快手科技	6500
4	滴滴出行	28000
5	百度公司	5500
6	美团点评	6000
7	腾讯公司	2000
8	阿里巴巴集团	5000

```
In [8]: #计算差额
diff = dfincome.merge(dfinvo,how='outer',left_on='客户名称',right_on='购方名称')
diff['差额'] = diff['凭证金额']-diff['金额']
diff
```

Out[8]: **客户名称 凭证金额 购方名称 金额 差额**

0	京东集团	10000.00	京东集团	2500	7500.00
1	华为集团	10000.00	华为集团	31000	-21000.00
2	NaN	NaN	小米公司	6000	NaN
3	NaN	NaN	快手科技	6500	NaN
4	滴滴出行	10000.00	滴滴出行	28000	-18000.00
5	NaN	NaN	百度公司	5500	NaN
6	NaN	NaN	美团点评	6000	NaN
7	腾讯公司	10000.00	腾讯公司	2000	8000.00
8	NaN	NaN	阿里巴巴集团	5000	NaN

In []:

In []:

In []: