In [14]:
```python
import pandas as pd
import numpy as np
import seaborn as sns
from matplotlib import pyplot as plt
from scipy.stats import chi2_contingency
from scipy import stats
```

In [2]:
```python
johny_talker = pd.read_excel("JohnyTalkers.xlsx")
johny_talker
```

Out[2]:

	Person	Drinks
0	Adults	Did Not Purchase
1	Adults	Did Not Purchase
2	Adults	Did Not Purchase
3	Adults	Did Not Purchase
4	Adults	Did Not Purchase
...	...	...
1215	Children	Did Not Purchase
1216	Children	Did Not Purchase
1217	Children	Did Not Purchase
1218	Children	Did Not Purchase
1219	Children	Purchased

1220 rows × 2 columns

In [4]:
```python
pd.get_dummies(johny_talker["Person"])
```

Out[4]:

	Adults	Children
0	1	0
1	1	0
2	1	0
3	1	0
4	1	0
...	...	...
1215	0	1
1216	0	1
1217	0	1
1218	0	1
1219	0	1

1220 rows × 2 columns

In [10]:
```python
person_data = pd.crosstab(index = johny_talker['Person'] , columns = johny_talker['D
person_data
```

Out[10]:

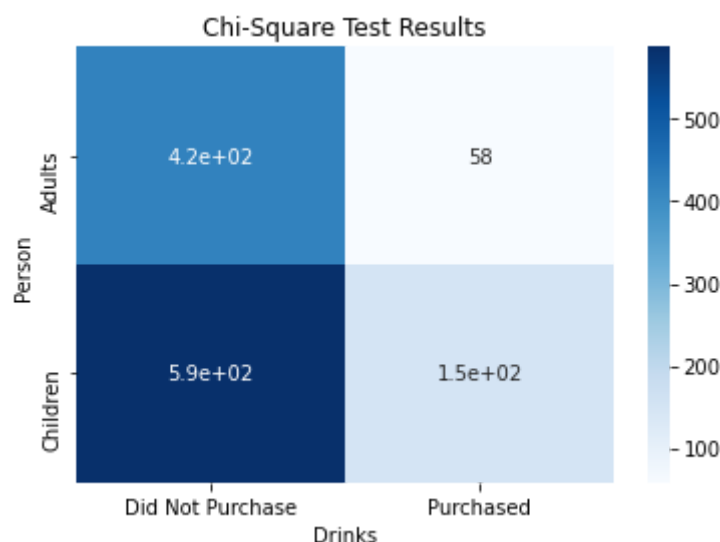| Drinks	Did Not Purchase	Purchased
Person		
Adults	422	58
Children	588	152

chi2-test

becouse in data set 2 categorical variables are given

In [13]:
```python
chi2_score,p_value,dof,expected_tabel = stats.chi2_contingency(observed = person_dat
print("chi2 - score        :"       , chi2_score)
print("p - value           :"       , p_value  )
print("degree of freedom  :"        , dof)
print("expected- tabel     :"        , expected_tabel )
```

```
chi2 - score       : 14.025770528307657
p - value          : 0.00018032227012050893
degree of freedom  : 1
expected- tabel    : [[397.37704918  82.62295082]
 [612.62295082 127.37704918]]
```

In [18]:
```python
sns.heatmap(data = person_data,annot = True,cmap = 'Blues')
plt.title("Chi-Square Test Results")
plt.show()
```


In [ ]: