PHD

PREDICTING THE DEMAND OF CONSUMABLE RETAIL PRODUCTS

Problem Statement:

A large company that manufactures eatables, delivers them through a group of suppliers to retailers of different kinds all over the country. One of the company's recent sustainable initiatives was to sell eatable products with no preservatives. These consumables have a shelf life of one week.

The company, to ensure the quality of its products on retailers' shelves, creates a program to buy back the remaining products on the shelves of the retailers post expiry (one week). This has significantly increased the company's operation costs and is also affecting its gross margins.

Observations:

- Based on data given, the train data size is: (53364883, 10).
- The test data size is: (20815581, 7).
- Creating a target variable by using the following command:

"train['demand_projection'] = train['num_units_sold_in_week']-train['num_units_returned']"

- Droping columns since train and test should contain same number of columns for the prediction process by using the following command:

"columnDrop = ['num_units_sold_in_week','sales_revenue_in_week', 'num_units_returned','returned_units_revenue_loss']

for a in columnDrop:

train.drop(a, axis = 1, inplace=True)

train.dtypes"

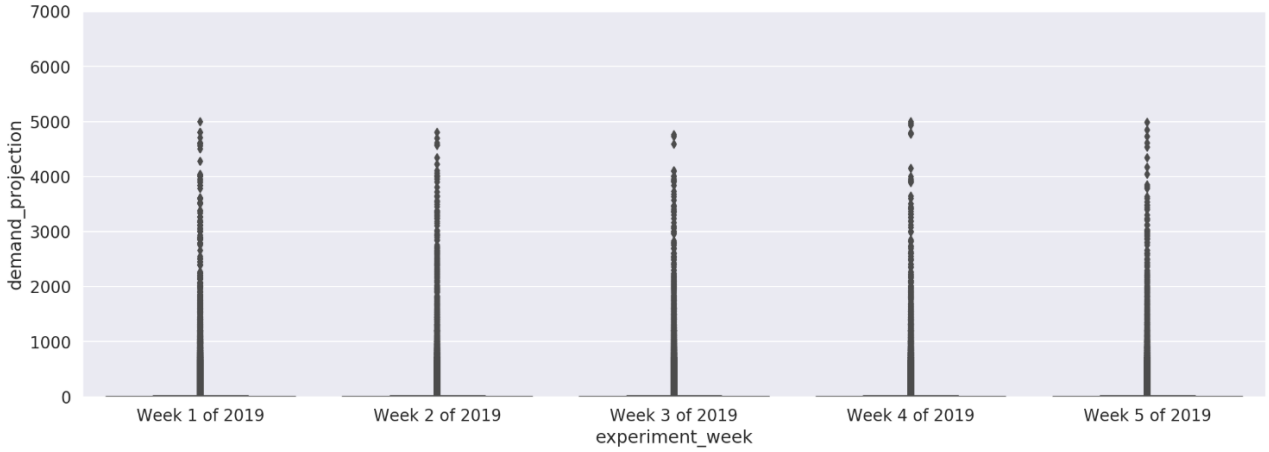
- Number of unique values in each column:

experiment_week	**5**
channel_type	**9**
num_units_sold_in_week	**1883**
sales_revenue_in_week	**69844**
num_units_returned	**497**

returned_units_revenue_loss	**13040**
store_identifier	**587773**
product_identifier	**1736**
category_of_route	**3490**
supplier_identifier	**552**
demand_projection	**2332**

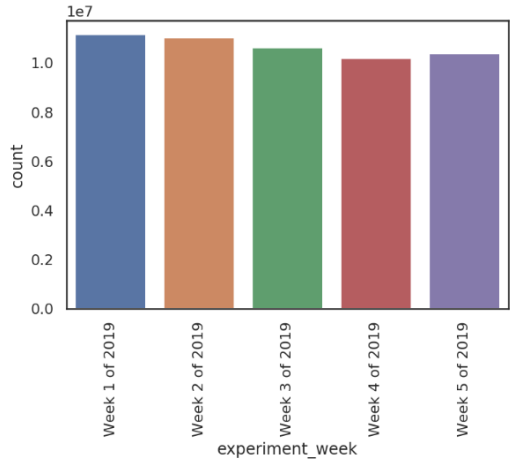
Data Visualization:

In [32]:
```
# experiment_week vs demand_projection
var = 'experiment_week'
data = pd.concat([train['demand_projection'], train[var]], axis=1)
f, ax = plt.subplots(figsize=(18, 6))
fig = sns.boxplot(x=var, y="demand_projection", data=data)
fig.axis(ymin=0, ymax=7000);
```

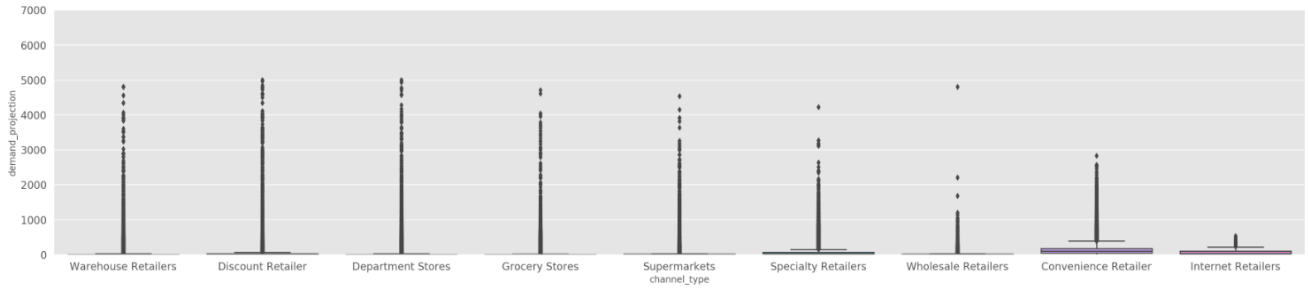


In [37]:
```
sns.countplot(train.experiment_week)
plt.xticks(rotation=90)
```

Out[37]: (array([0, 1, 2, 3, 4]), <a list of 5 Text xticklabel objects>)

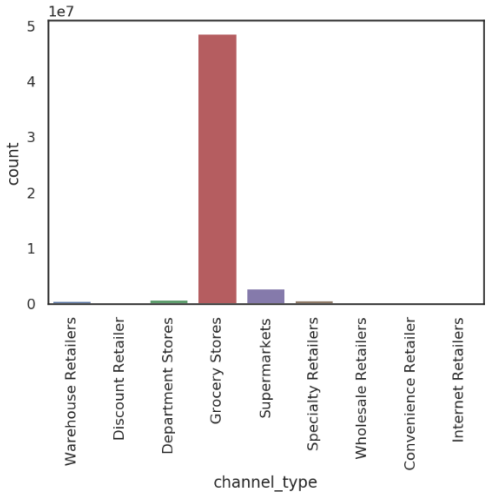


In [40]:
```
# channel_type vs demand_projection
var = 'channel_type'
data = pd.concat([train['demand_projection'], train[var]], axis=1)
f, ax = plt.subplots(figsize=(30, 6))
fig = sns.boxplot(x=var, y="demand_projection", data=data)
fig.axis(ymin=0, ymax=7000);
```



In [36]:
```
sns.countplot(train.channel_type)
plt.xticks(rotation=90)
```

Out[36]: (array([0, 1, 2, 3, 4, 5, 6, 7, 8]), <a list of 9 Text xticklabel objects>)



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

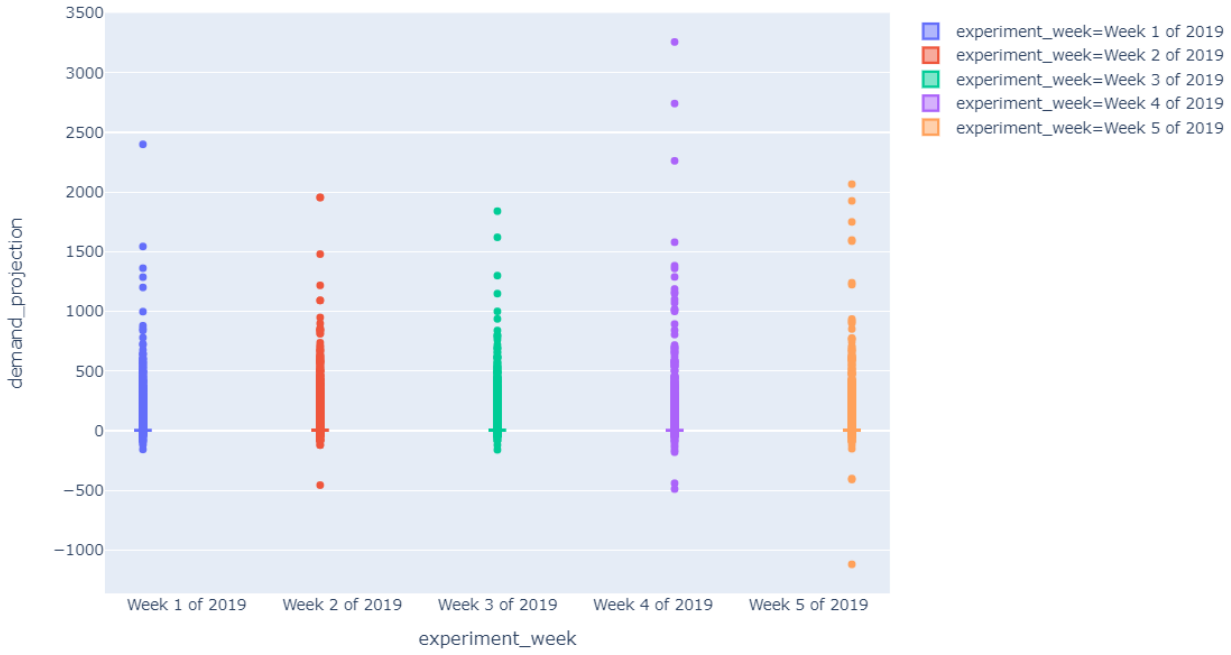
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [7]:
```
import plotly.express as px
px.box(train, x="experiment_week", y="demand_projection", color="experiment_week")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

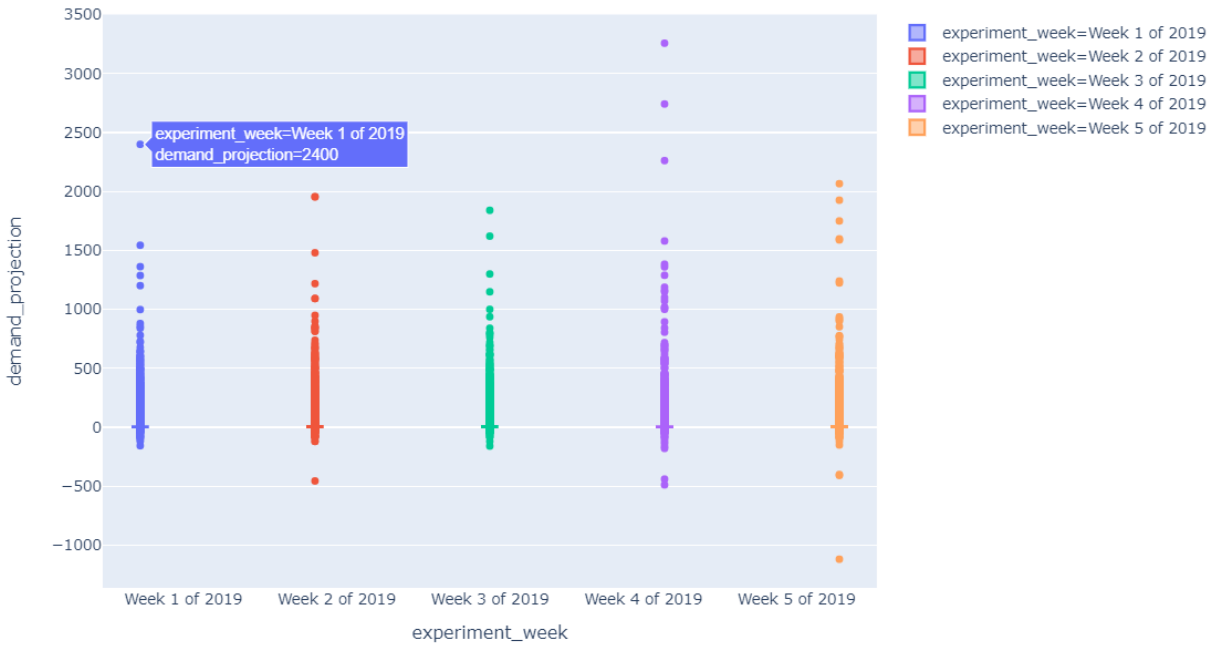
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [10]:
```
import plotly.express as px
px.box(train, x="experiment_week", y="demand_projection", color="experiment_week")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

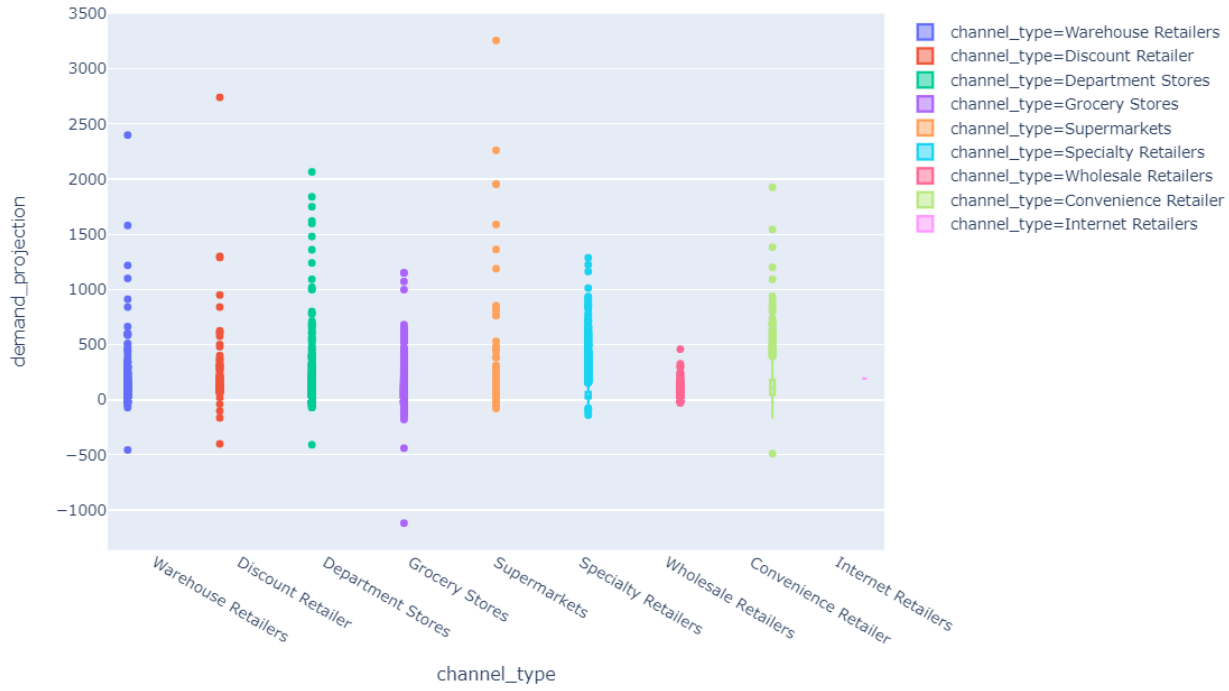
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [8]:
```
import plotly.express as px
px.box(train, x="channel_type", y="demand_projection", color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

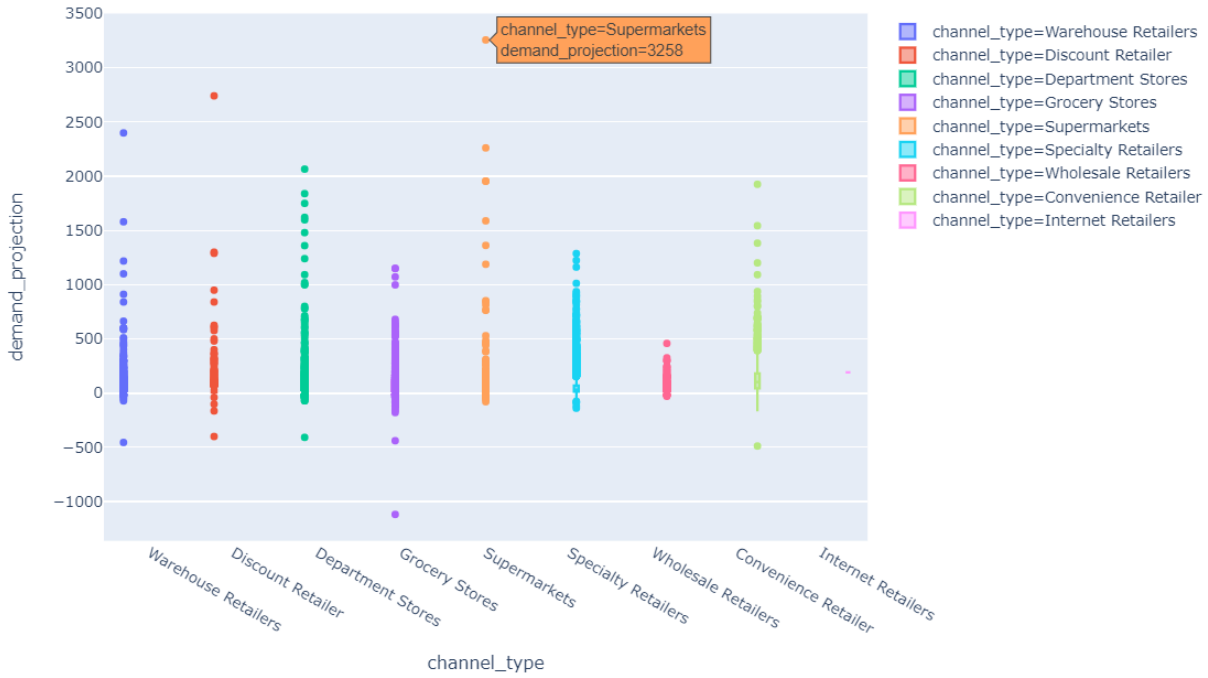
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [11]:
```
import plotly.express as px
px.box(train, x="channel_type", y="demand_projection", color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

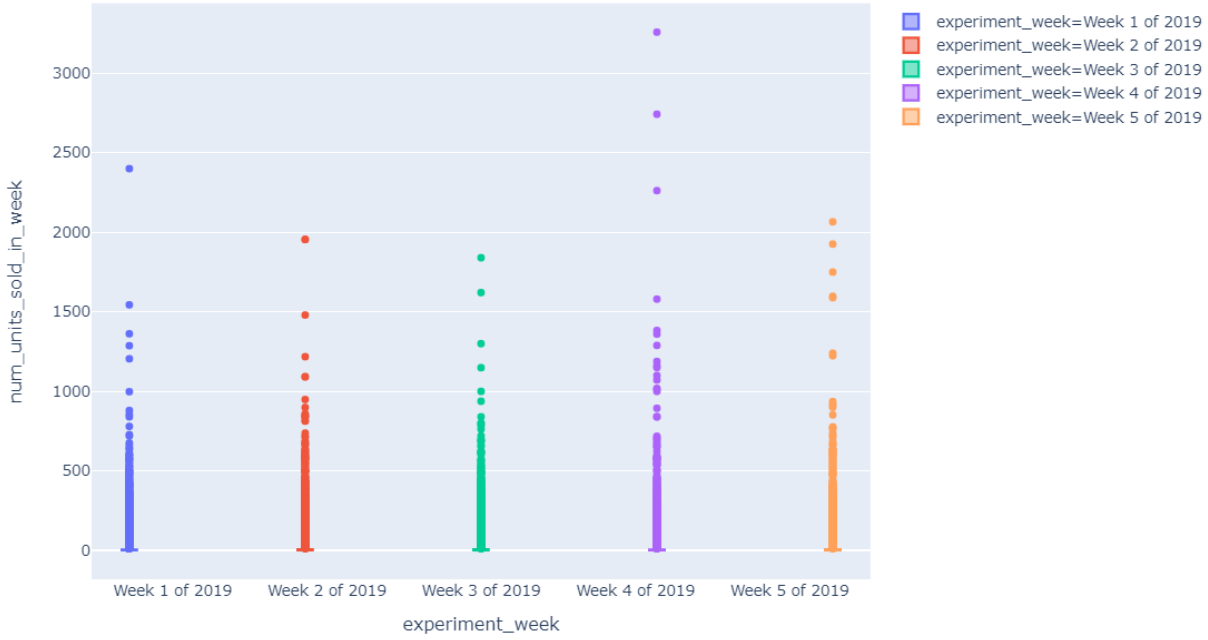
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [9]:
```
import plotly.express as px
px.box(train, x="experiment_week", y="num_units_sold_in_week", color="experiment_week")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

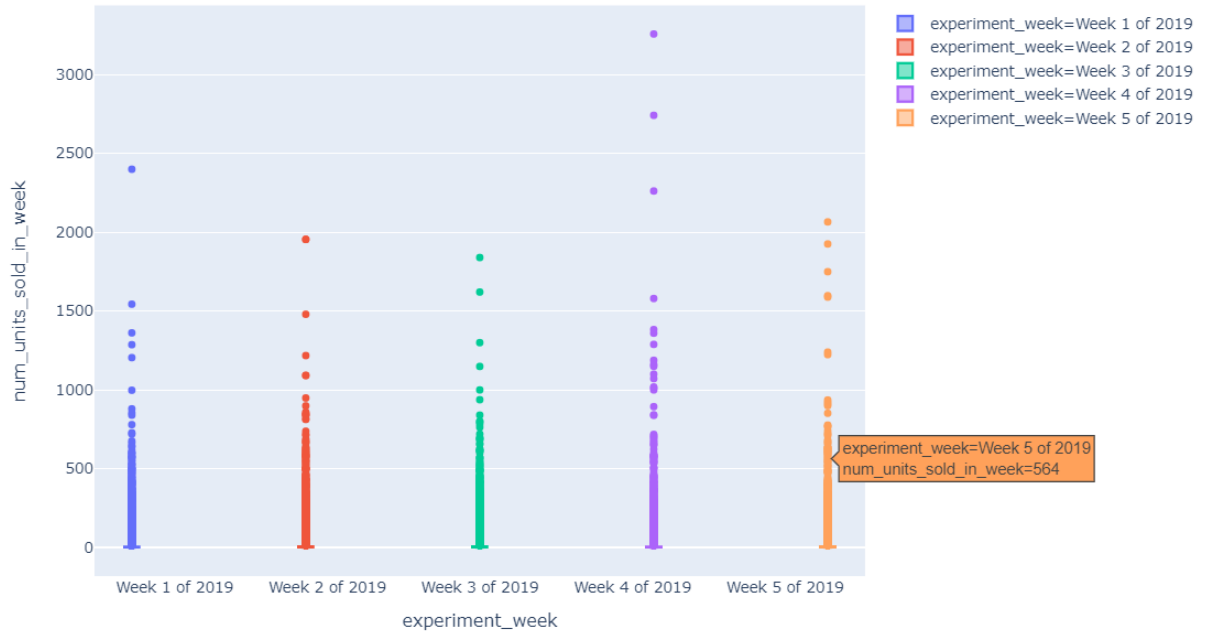
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [9]:
```
import plotly.express as px
px.box(train, x="experiment_week", y="num_units_sold_in_week", color="experiment_week")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

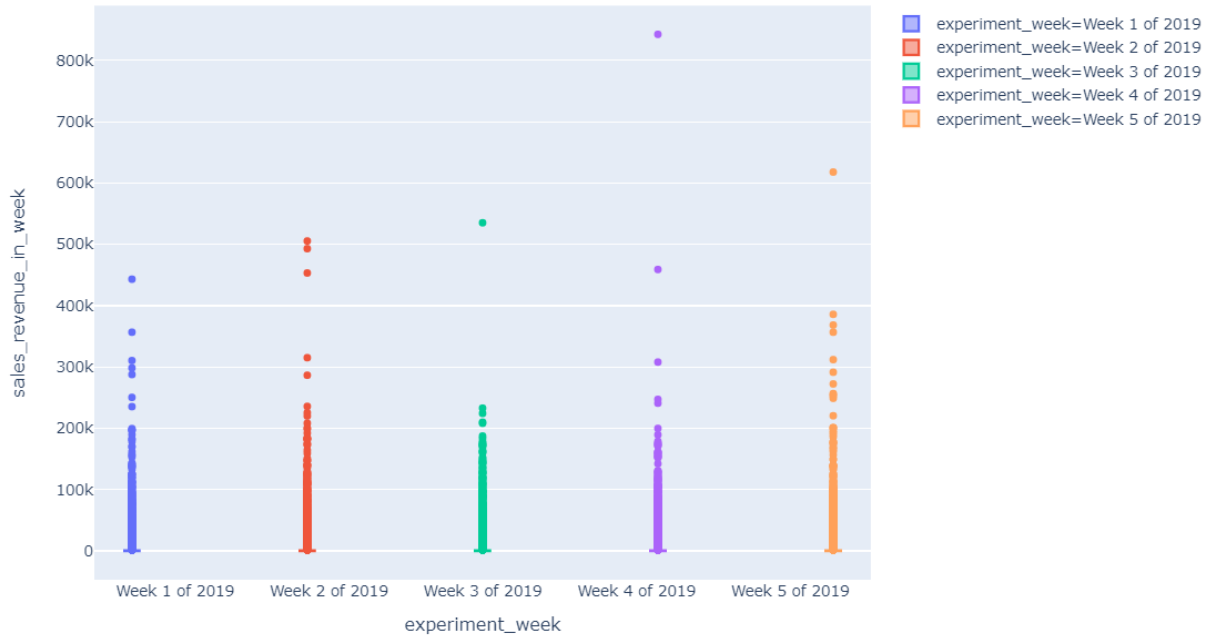
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [12]:
```
import plotly.express as px
px.box(train, x="experiment_week", y="sales_revenue_in_week", color="experiment_week")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

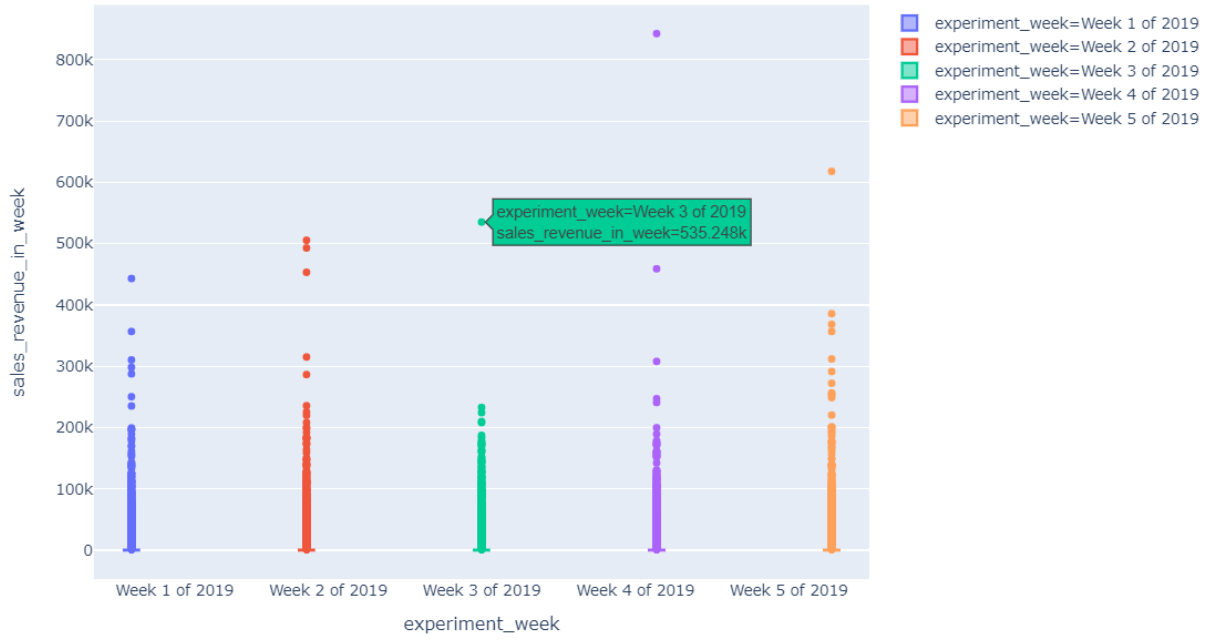
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [12]:
```
import plotly.express as px
px.box(train, x="experiment_week", y="sales_revenue_in_week", color="experiment_week")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

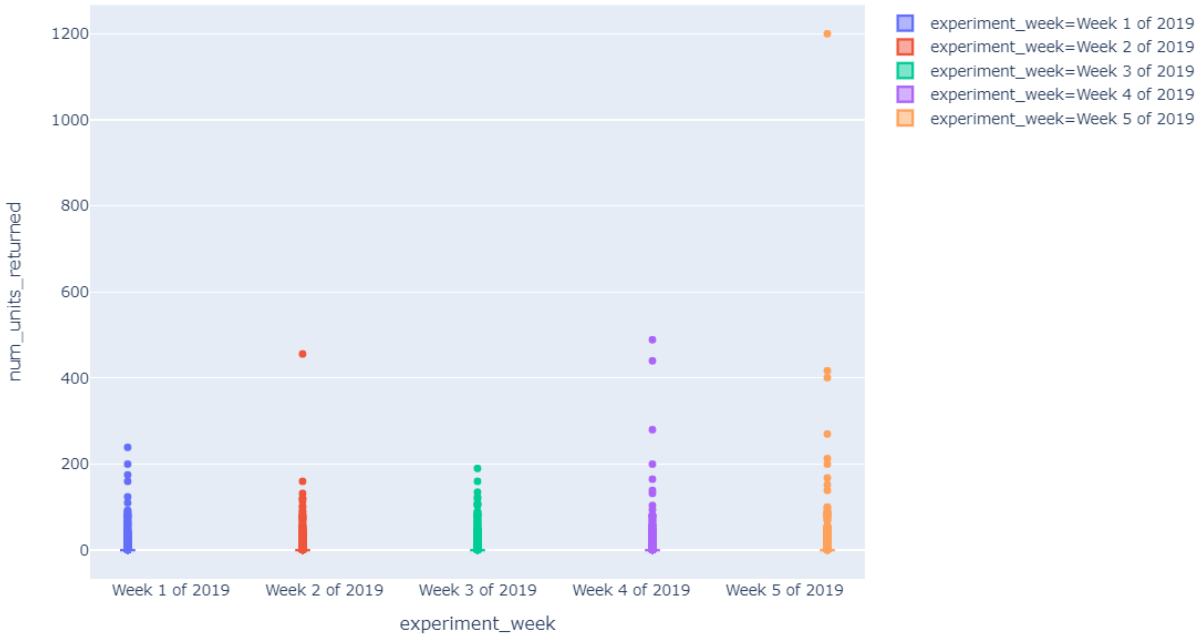
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [13]:
```
import plotly.express as px
px.box(train, x="experiment_week", y="num_units_returned", color="experiment_week")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

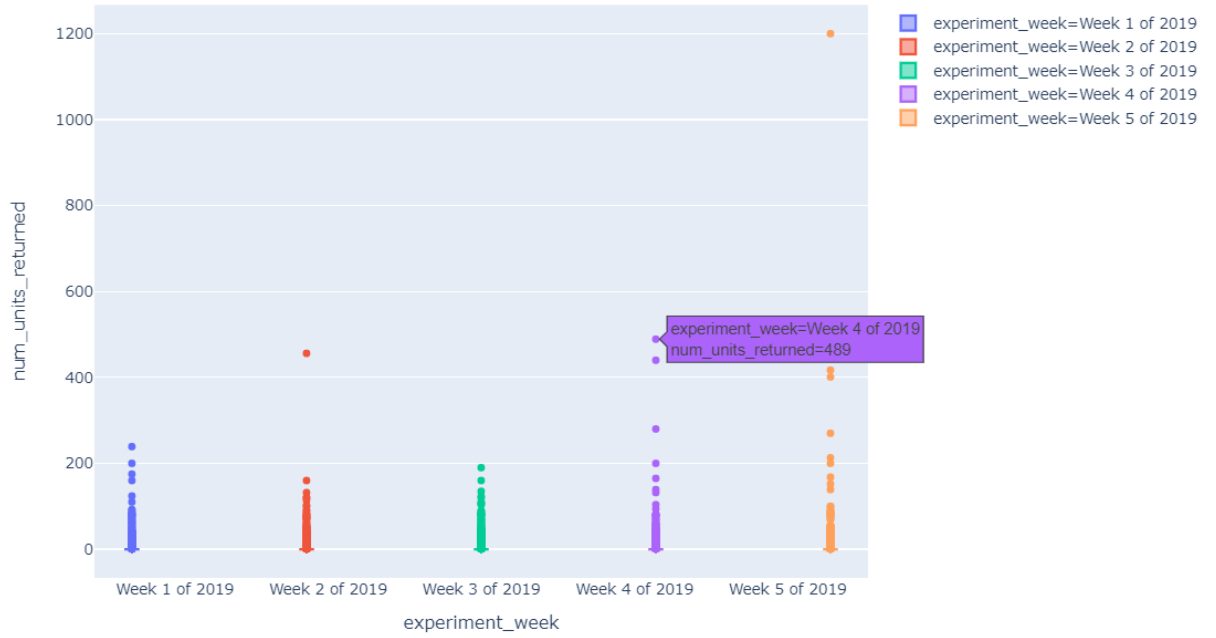
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [13]:
```
import plotly.express as px
px.box(train, x="experiment_week", y="num_units_returned", color="experiment_week")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

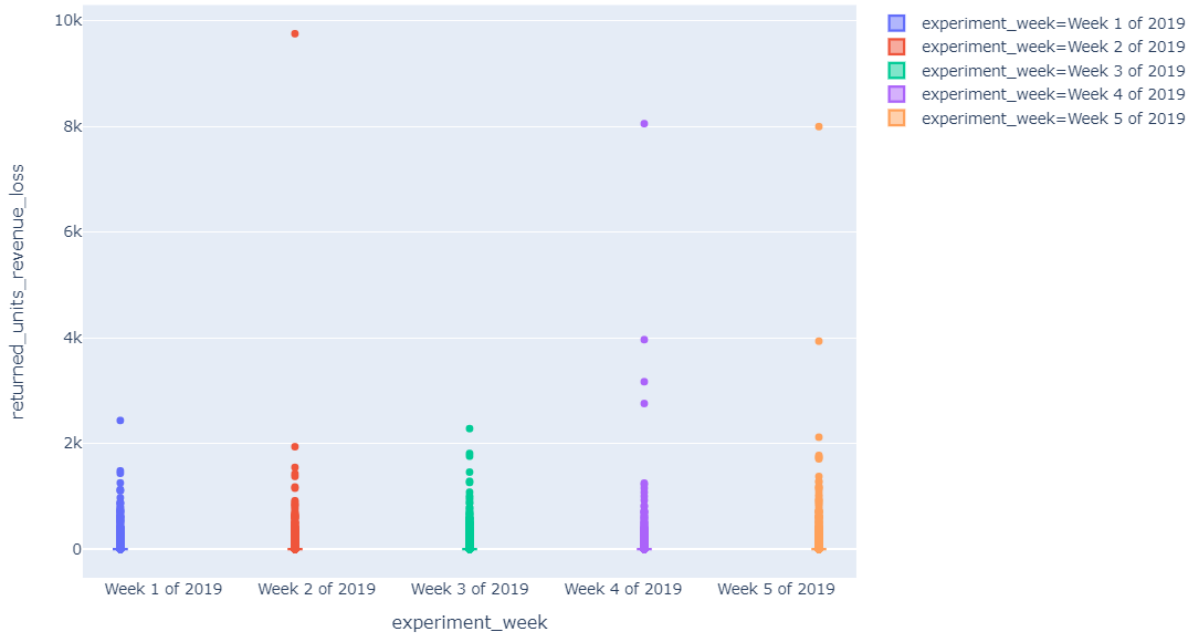
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [14]:
```
import plotly.express as px
px.box(train, x="experiment_week", y="returned_units_revenue_loss", color="experiment_week")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

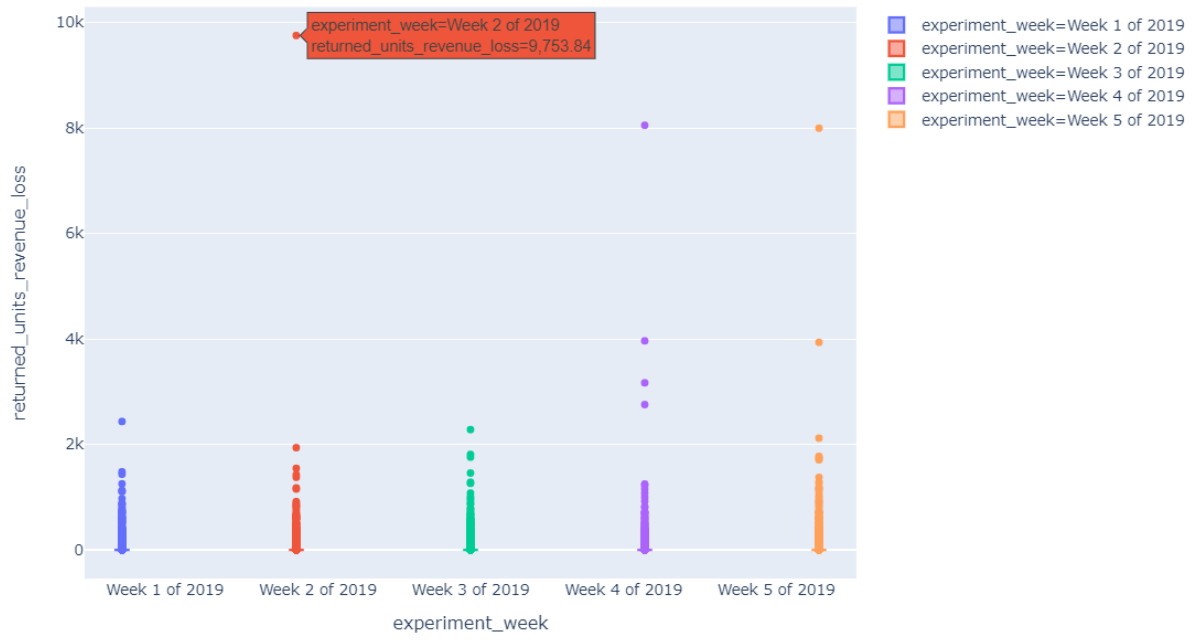
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [14]:
```
import plotly.express as px
px.box(train, x="experiment_week", y="returned_units_revenue_loss", color="experiment_week")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

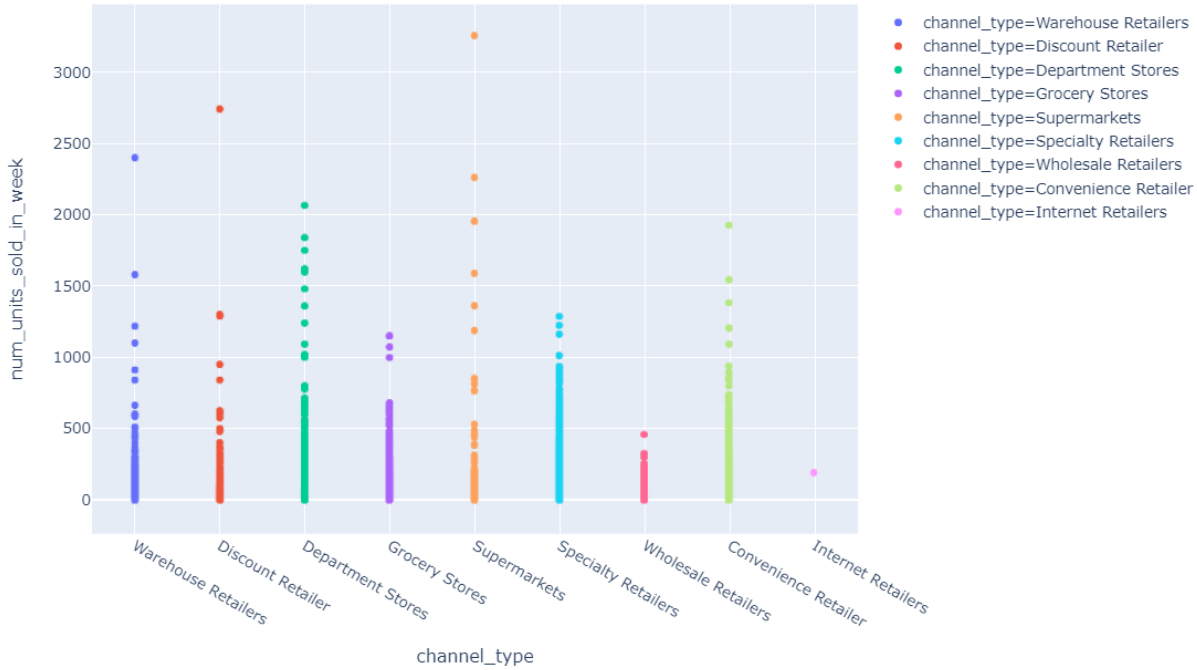
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [15]:
```
import plotly.express as px
px.scatter(train, x="channel_type", y="num_units_sold_in_week", color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

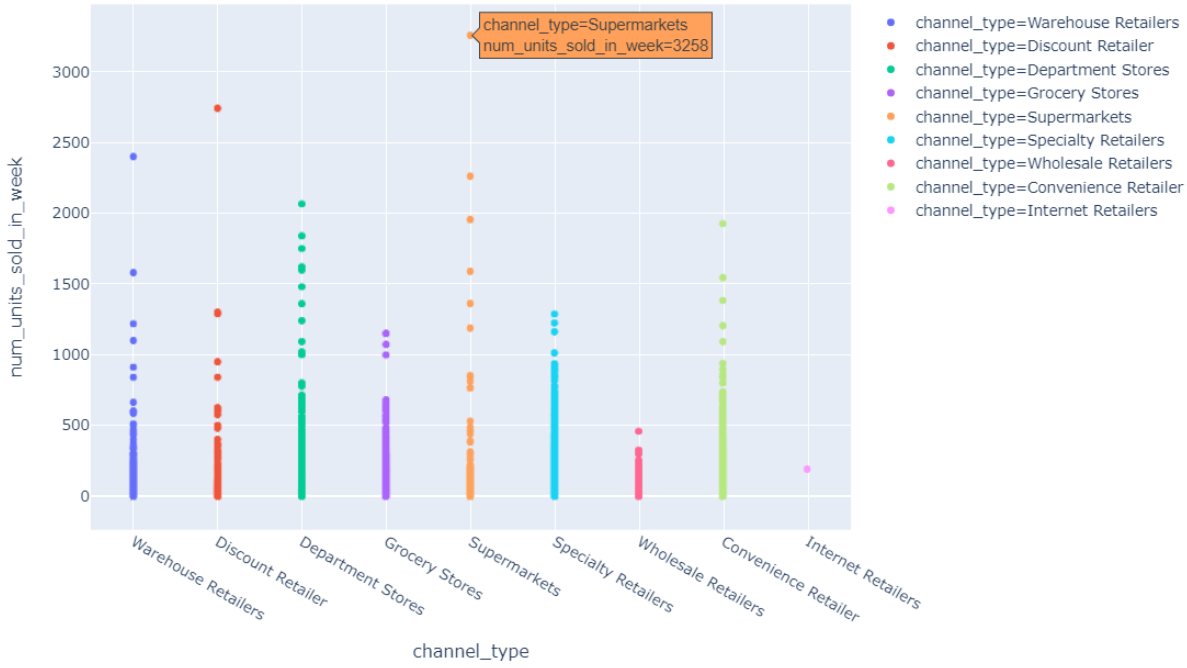
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [15]:
```
import plotly.express as px
px.scatter(train, x="channel_type", y="num_units_sold_in_week", color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

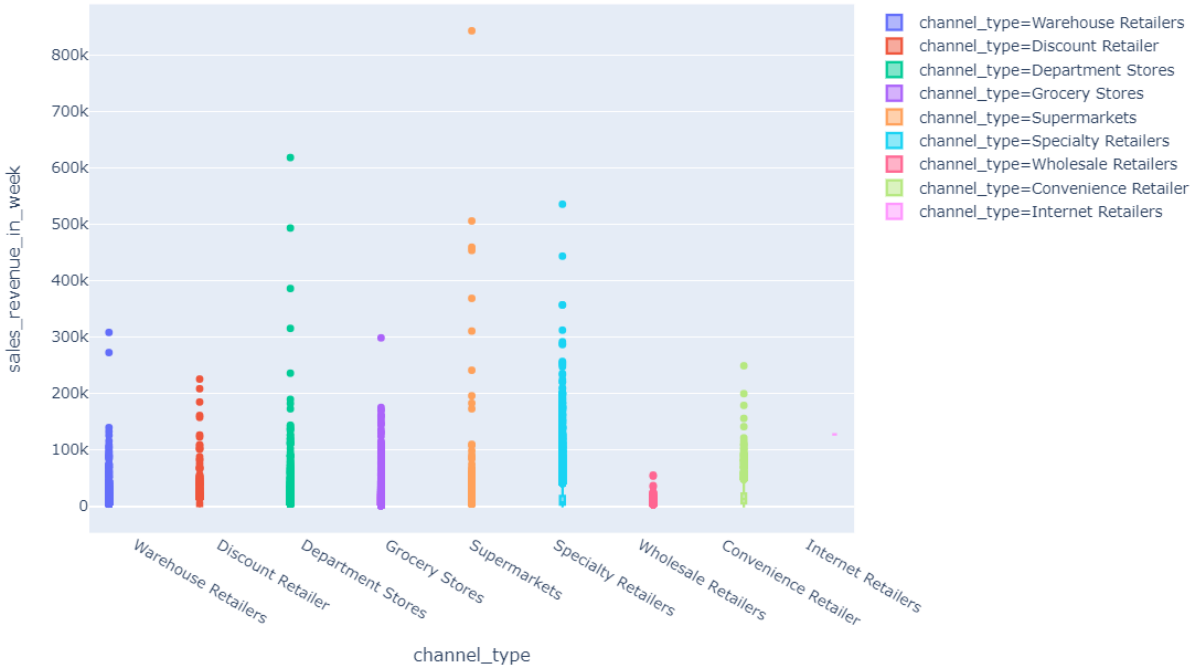
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [16]:
```
import plotly.express as px
px.box(train, x="channel_type", y="sales_revenue_in_week", color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

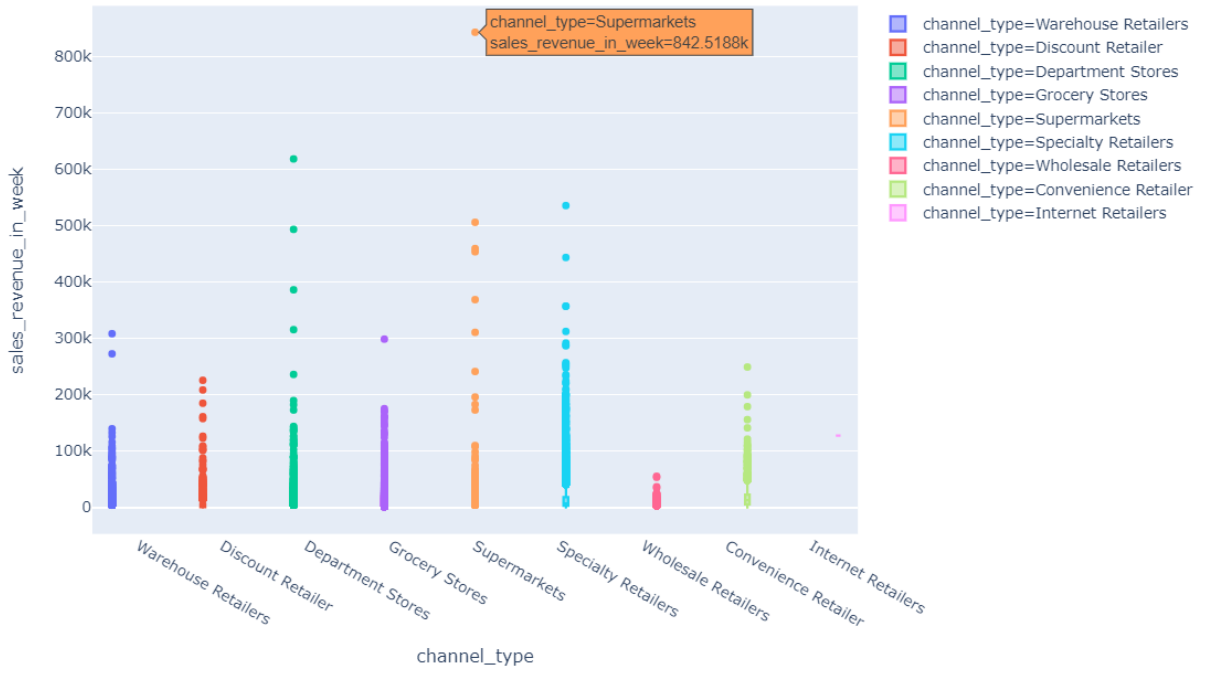
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [16]:
```
import plotly.express as px
px.box(train, x="channel_type", y="sales_revenue_in_week", color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

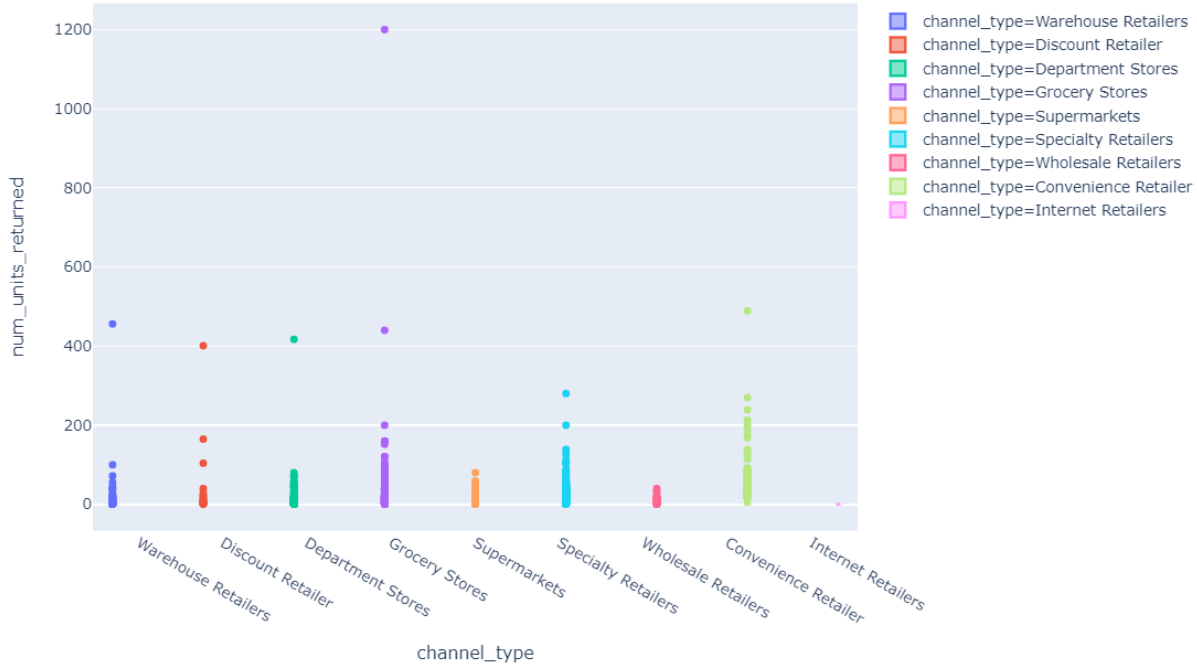
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [17]:
```
import plotly.express as px
px.box(train, x="channel_type", y="num_units_returned", color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

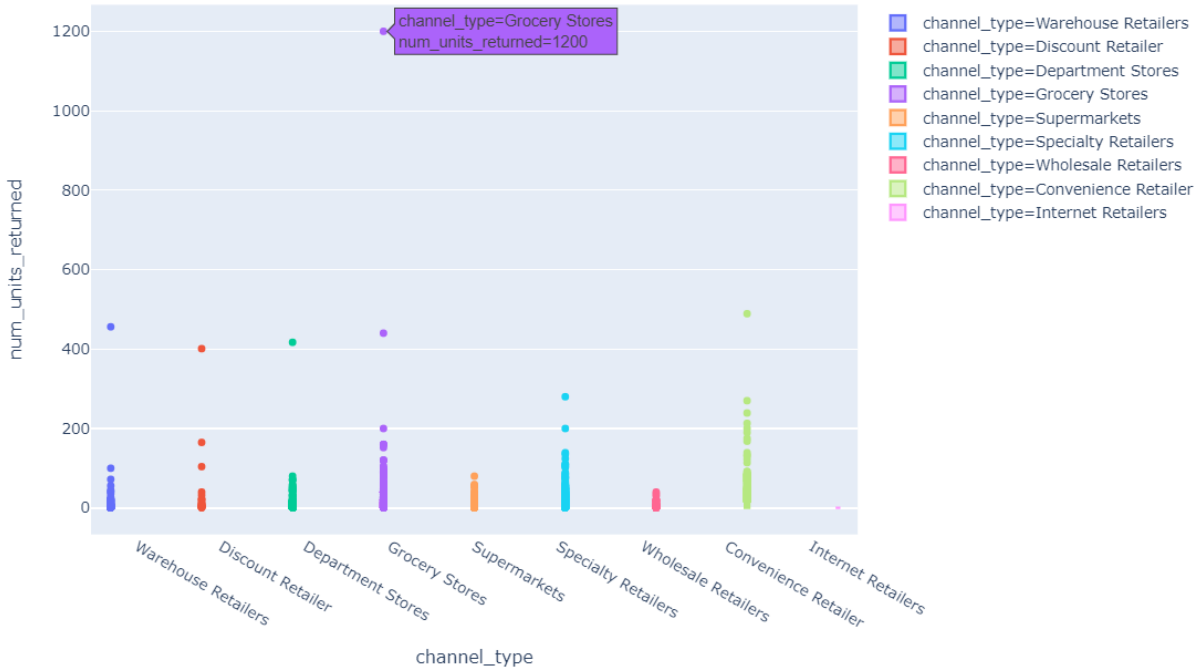
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [17]:
```
import plotly.express as px
px.box(train, x="channel_type", y="num_units_returned", color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

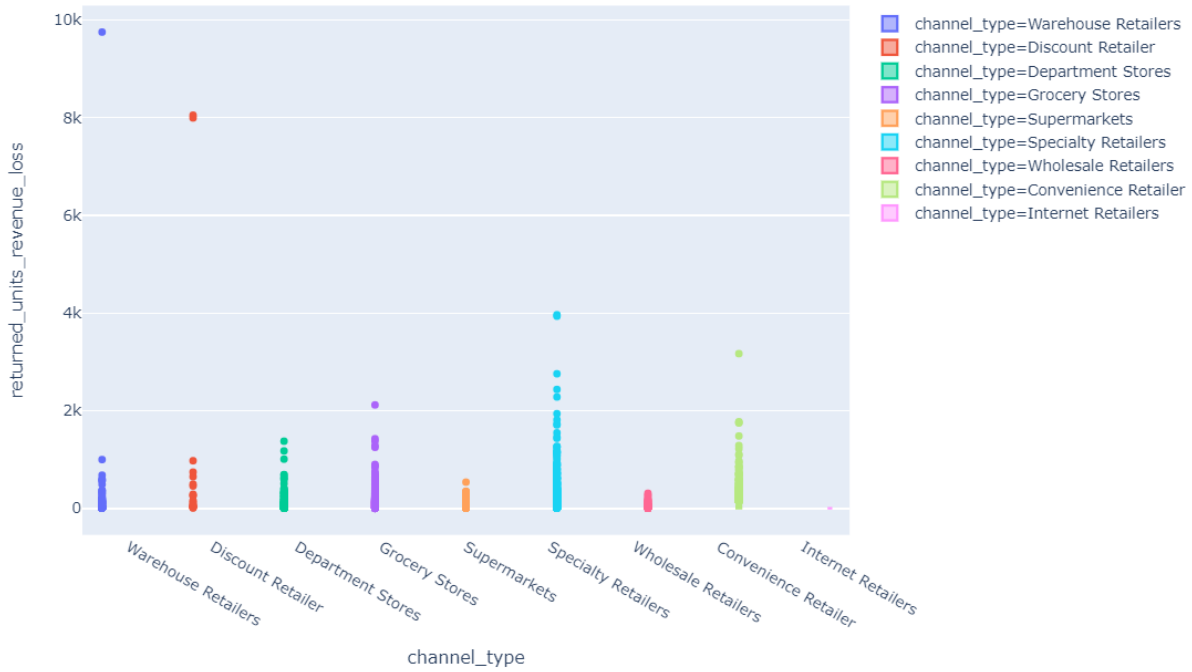
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [18]:
```
import plotly.express as px
px.box(train, x="channel_type", y="returned_units_revenue_loss", color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

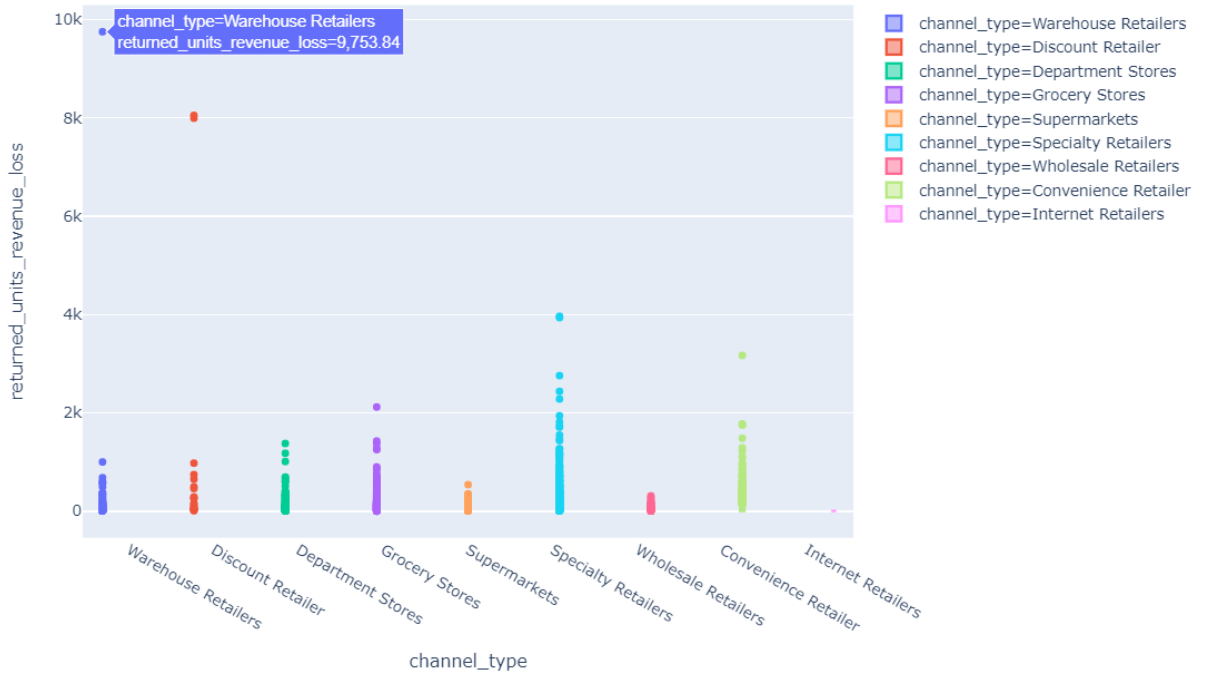
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [18]:
```
import plotly.express as px
px.box(train, x="channel_type", y="returned_units_revenue_loss", color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

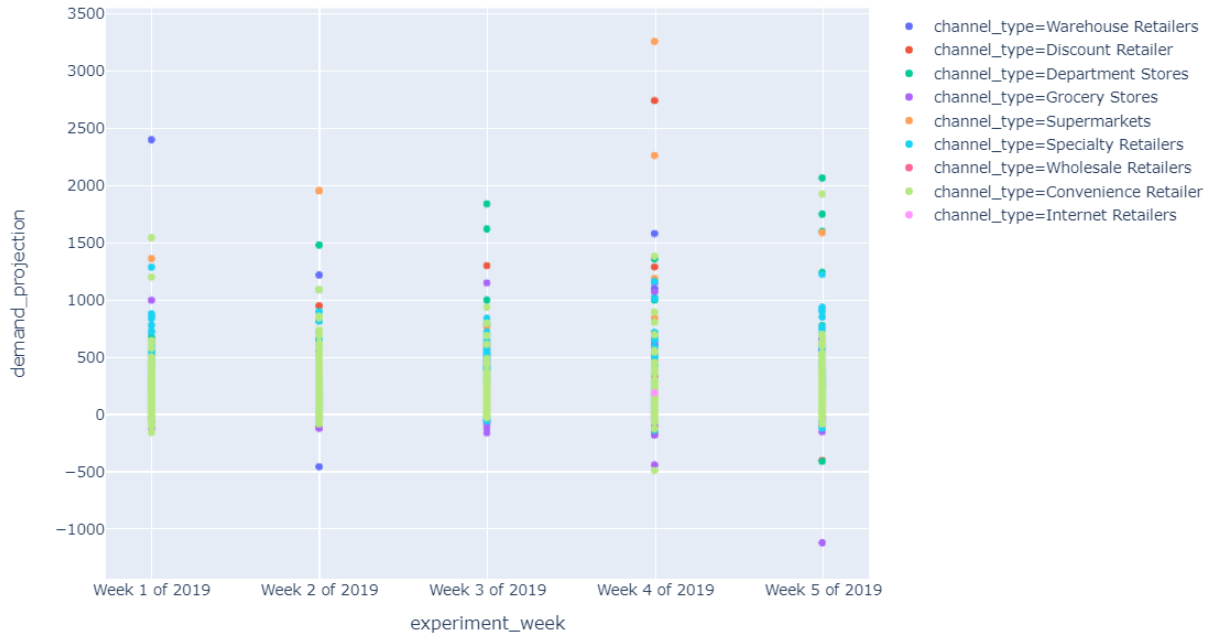
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [19]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="demand_projection", color="channel_type", size_max= 5)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

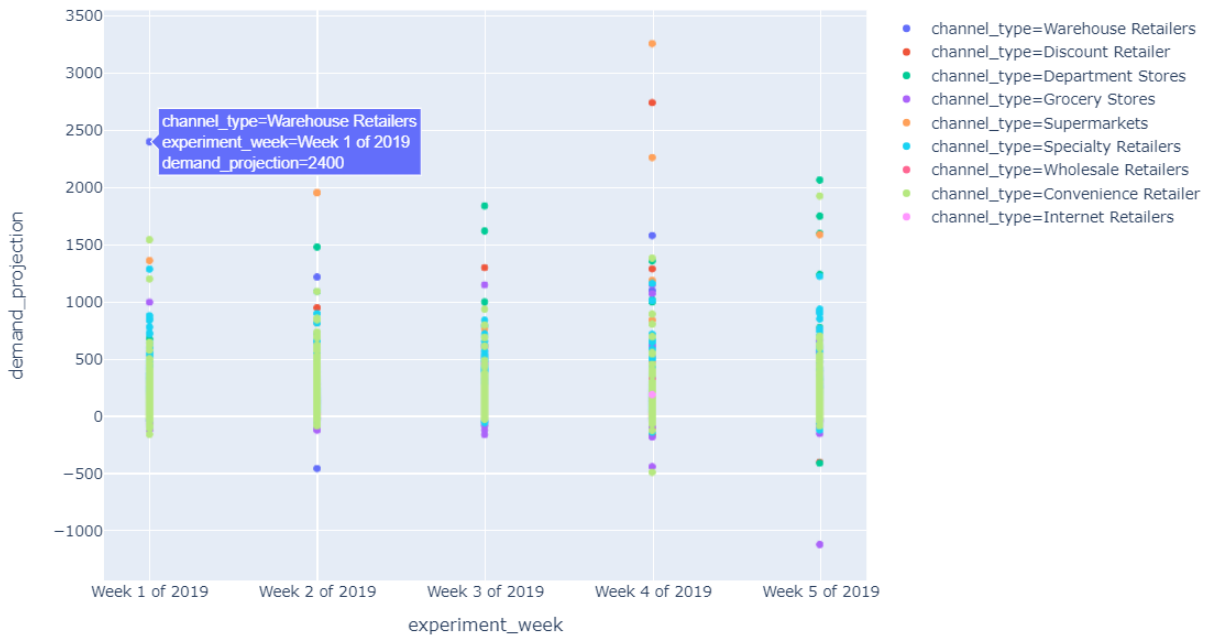
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [19]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="demand_projection", color="channel_type", size_max= 5)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

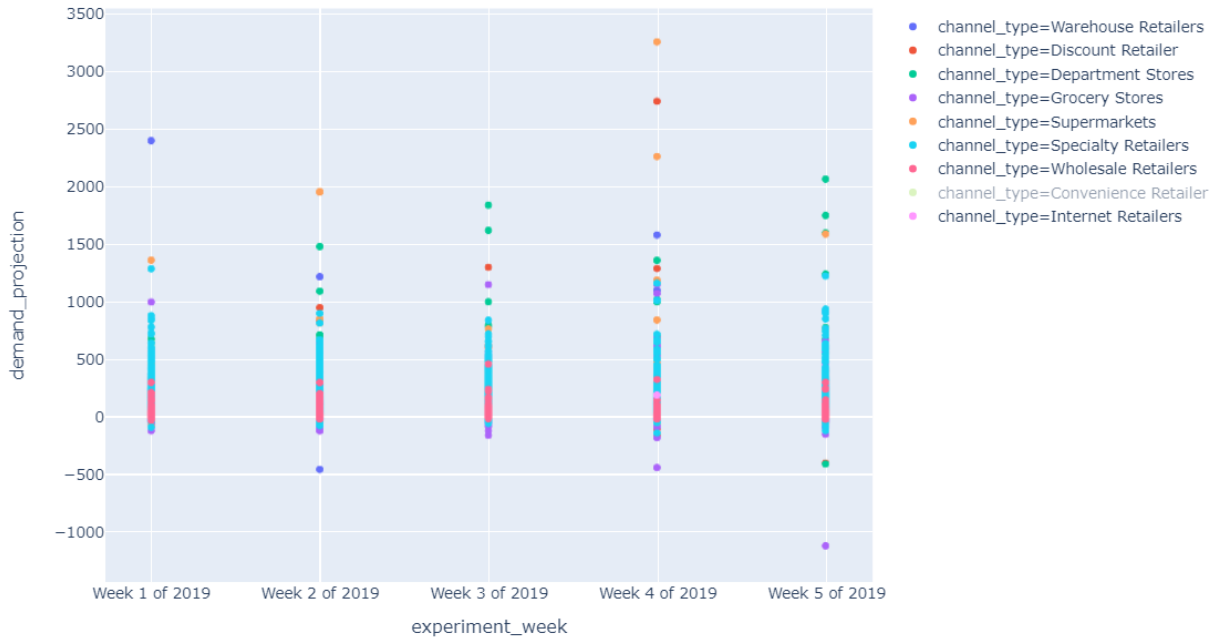
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [19]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="demand_projection", color="channel_type", size_max= 5)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

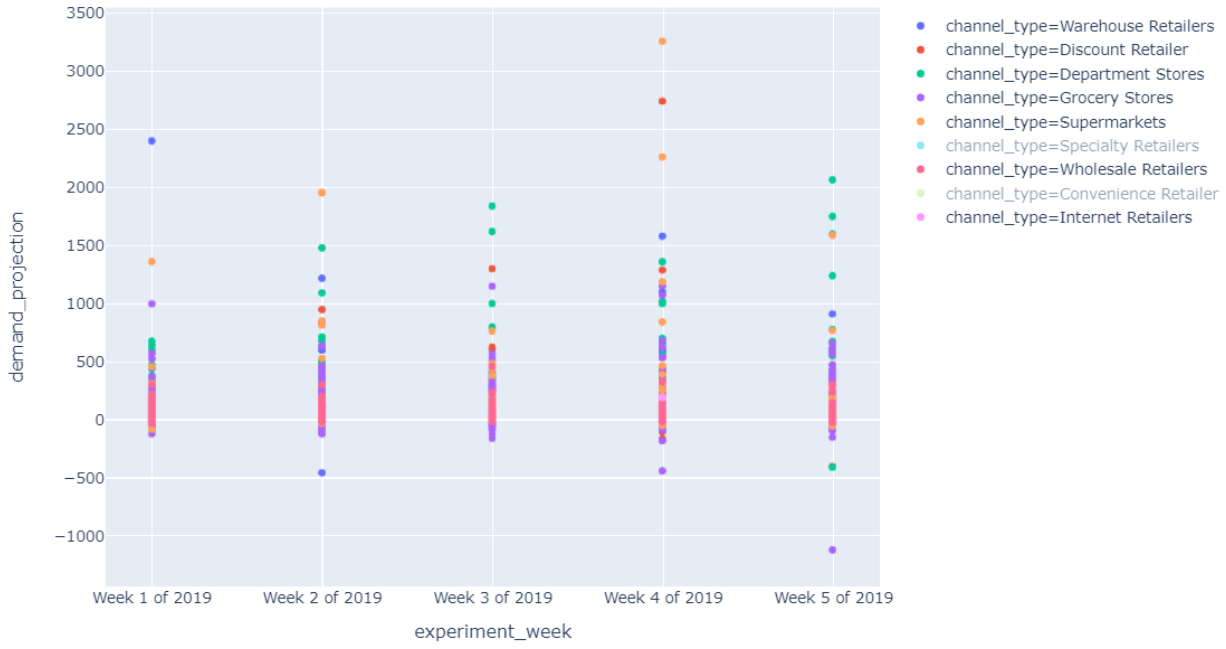
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [19]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="demand_projection", color="channel_type", size_max= 5)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

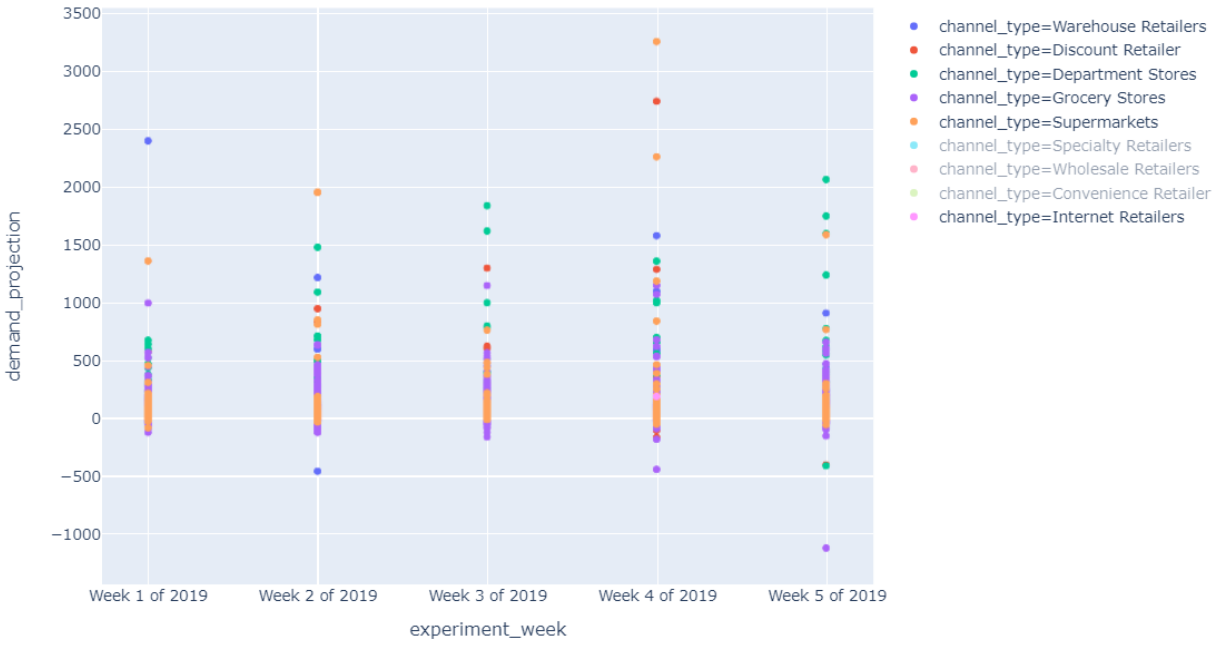
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [19]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="demand_projection", color="channel_type", size_max= 5)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

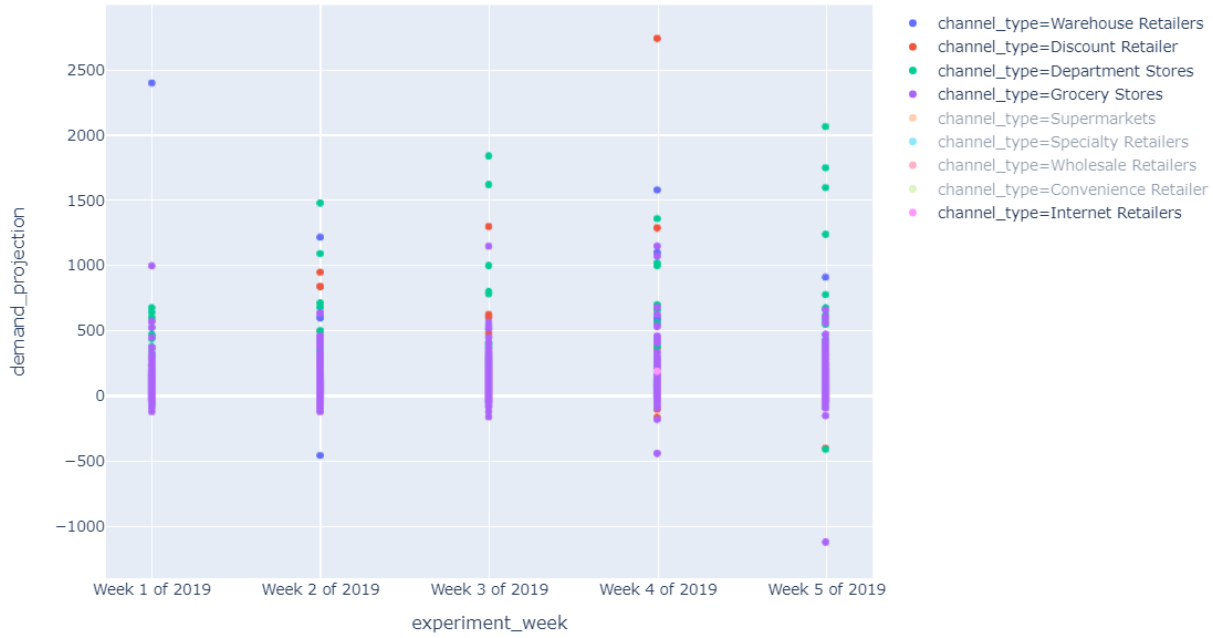
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [19]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="demand_projection", color="channel_type", size_max= 5)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

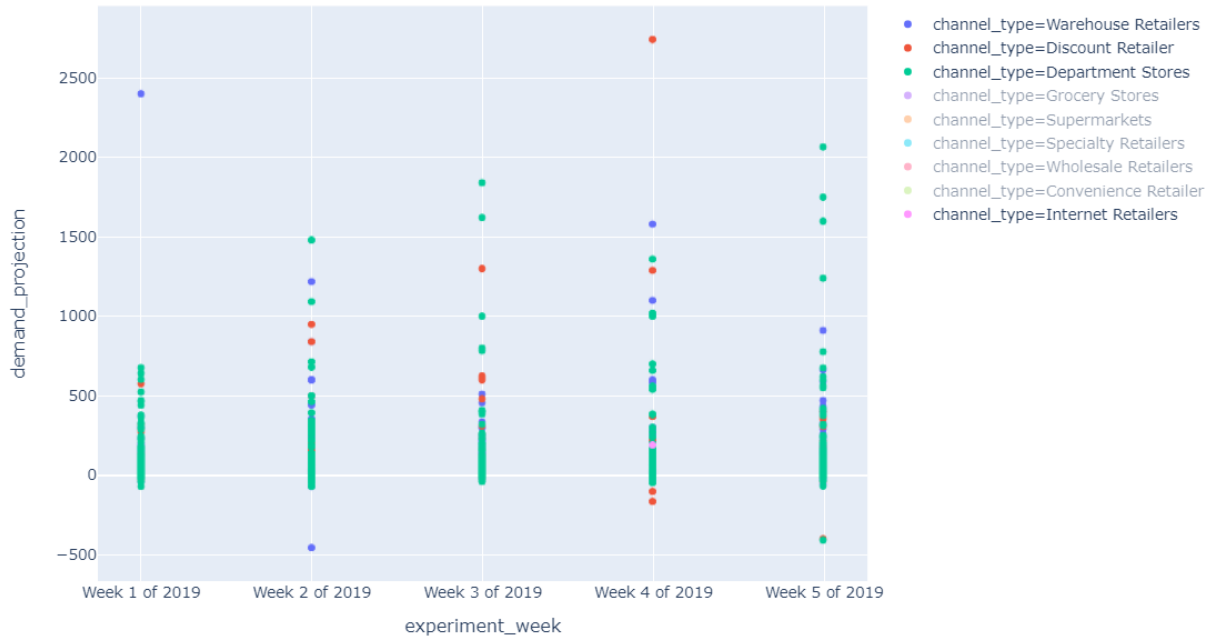
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [19]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="demand_projection", color="channel_type", size_max= 5)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

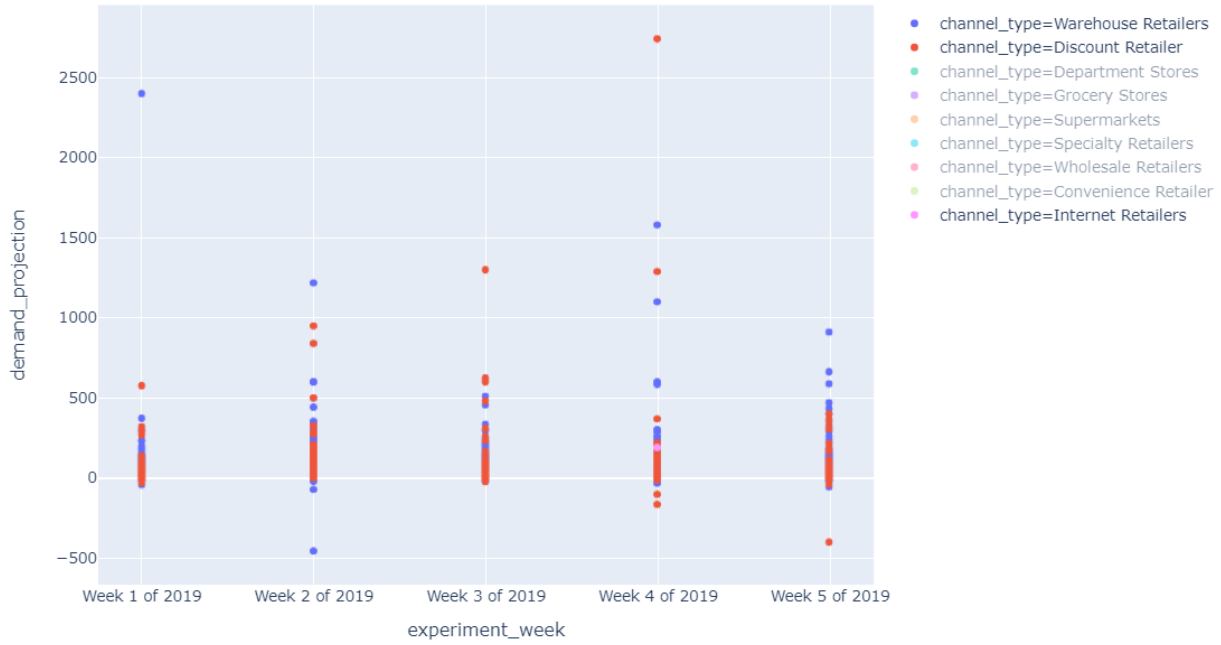
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [19]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="demand_projection", color="channel_type", size_max= 5)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

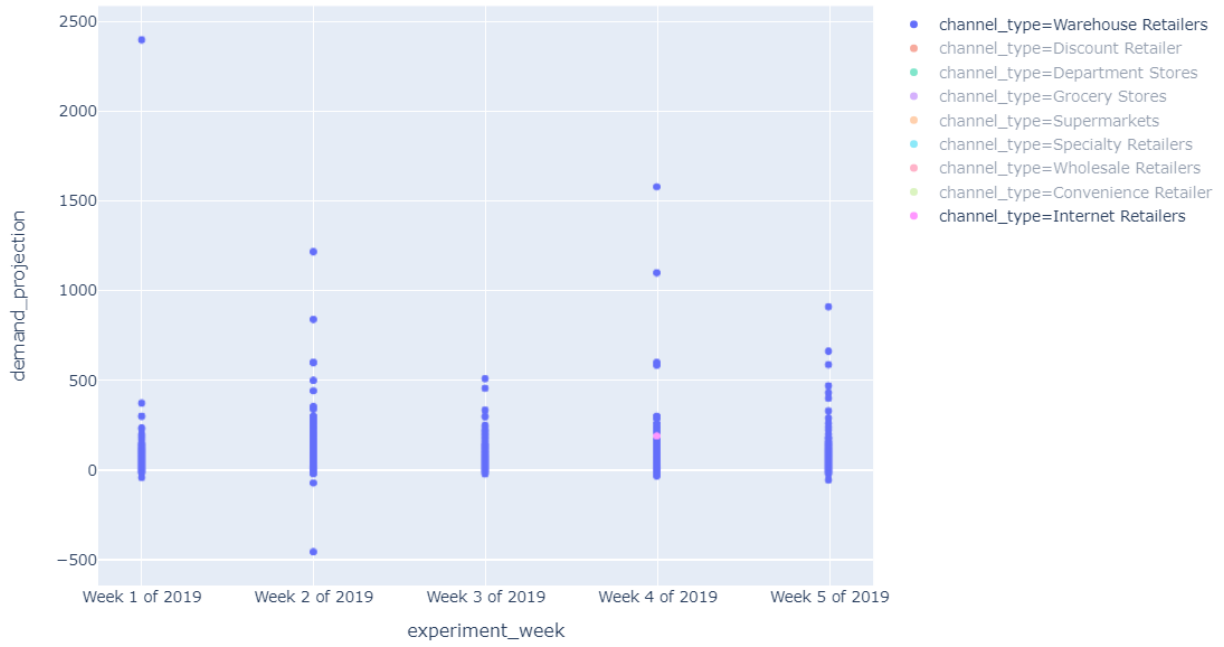
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [19]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="demand_projection", color="channel_type", size_max= 5)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

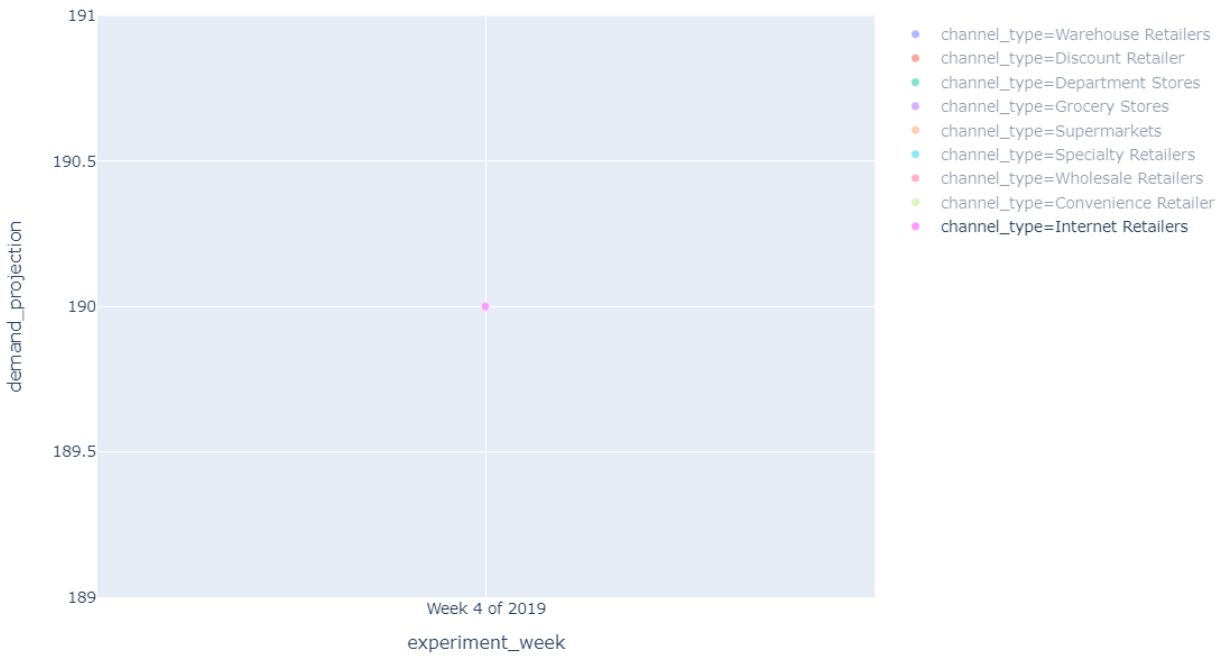
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [19]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="demand_projection", color="channel_type", size_max= 5)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [19]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="demand_projection", color="channel_type", size_max= 5)
```

jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

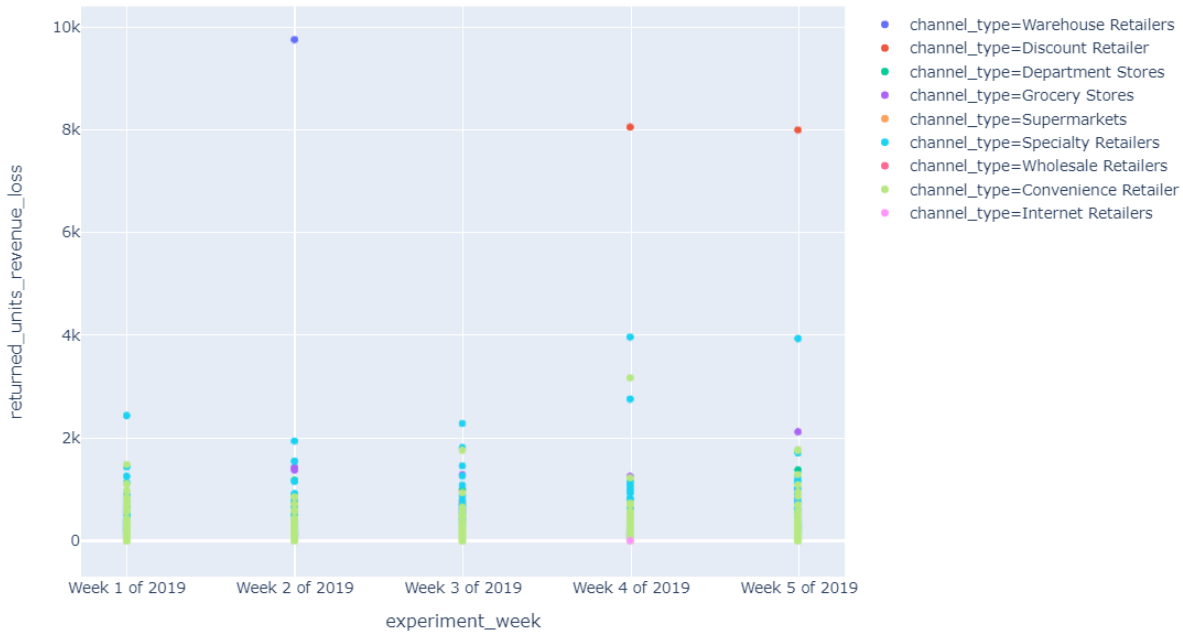
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [20]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="returned_units_revenue_loss", color="channel_type", size_max= 50)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

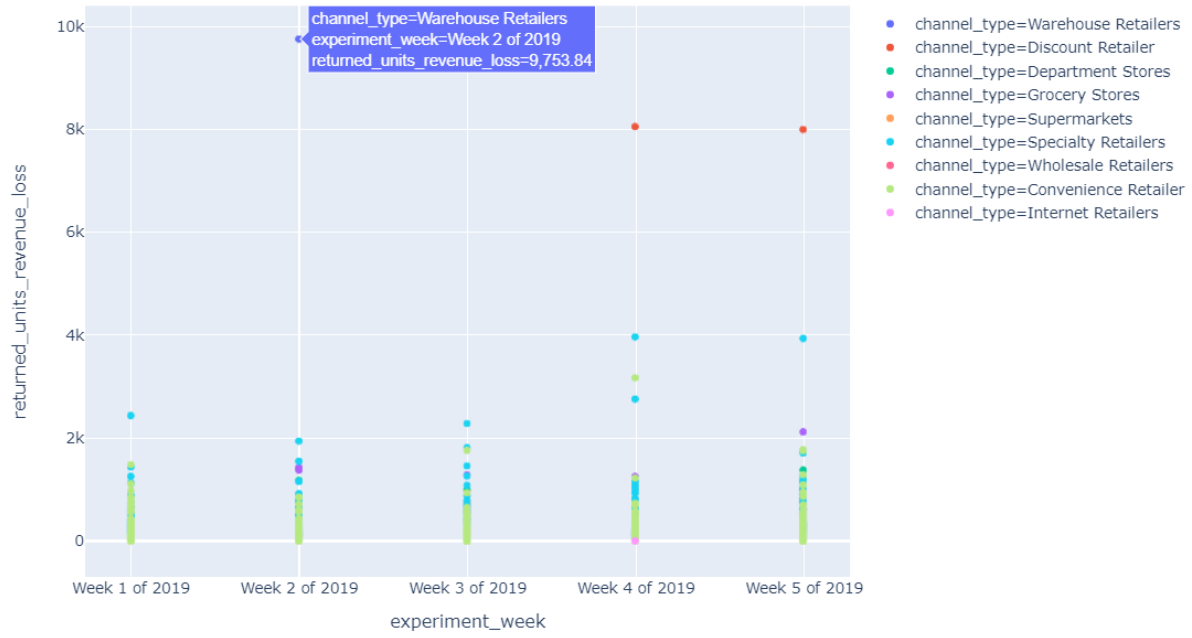
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [20]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="returned_units_revenue_loss", color="channel_type", size_max= 50)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

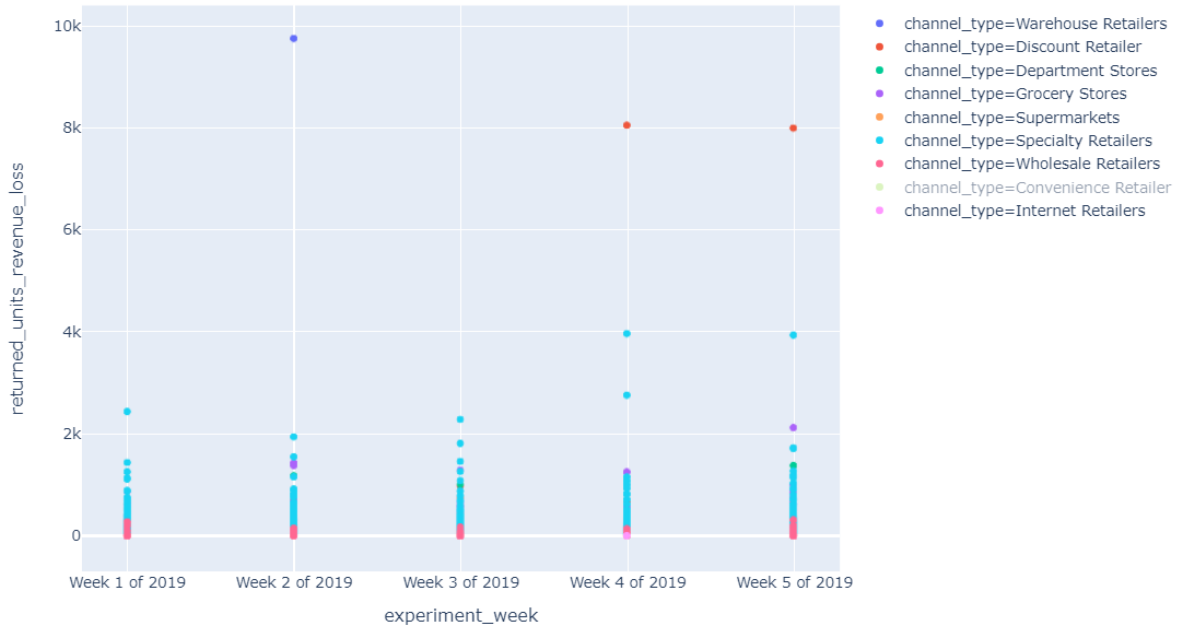
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [20]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="returned_units_revenue_loss", color="channel_type", size_max= 50)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

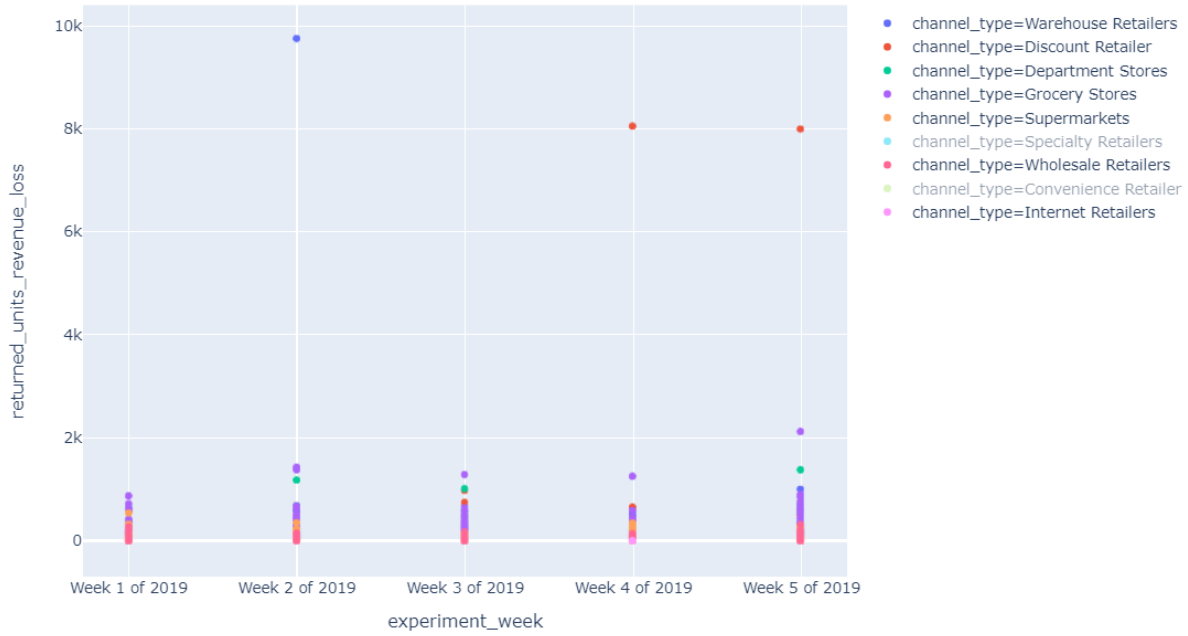
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [20]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="returned_units_revenue_loss", color="channel_type", size_max= 50)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

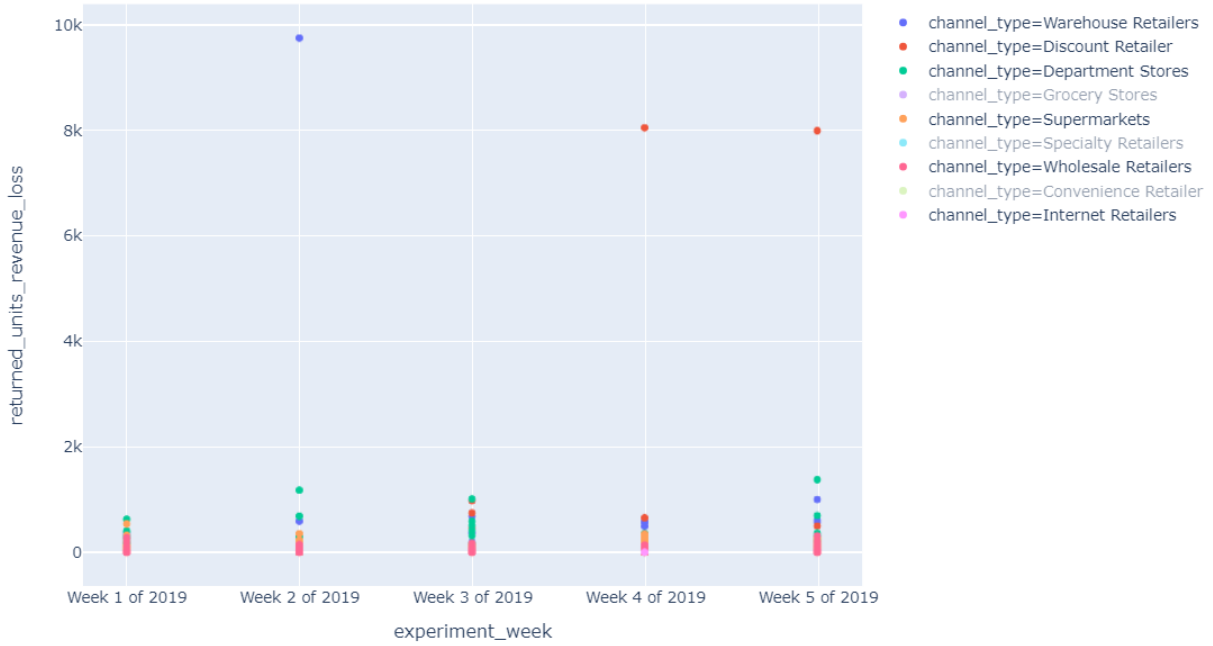
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [20]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="returned_units_revenue_loss", color="channel_type", size_max= 50)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

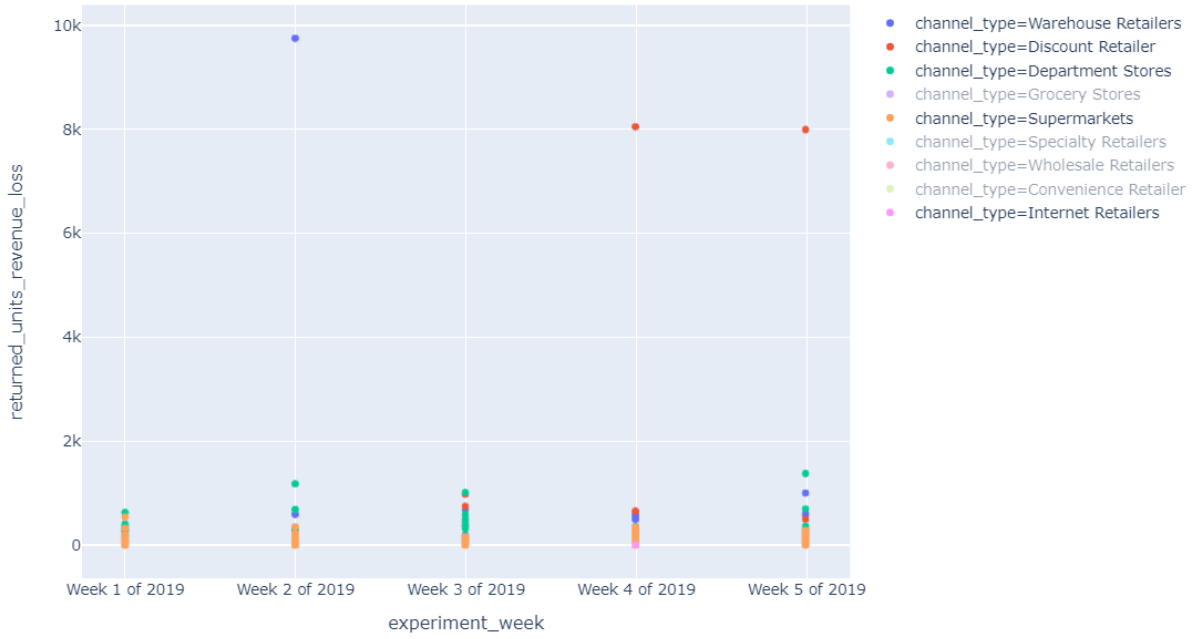
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [20]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="returned_units_revenue_loss", color="channel_type", size_max= 50)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

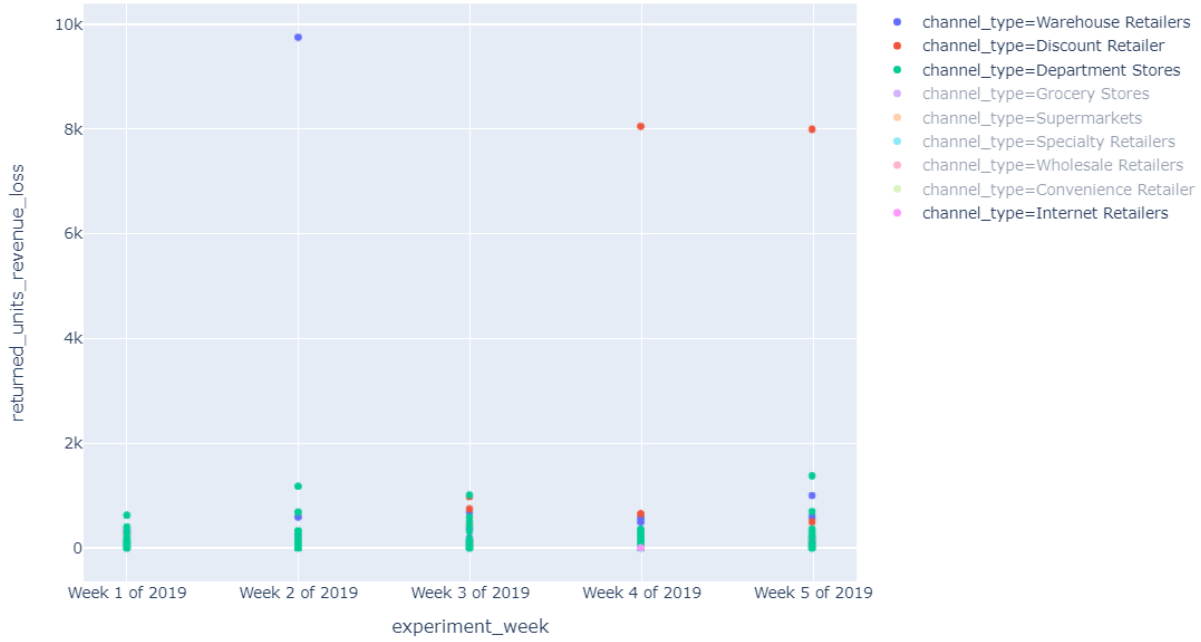
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [20]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="returned_units_revenue_loss", color="channel_type", size_max= 50)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

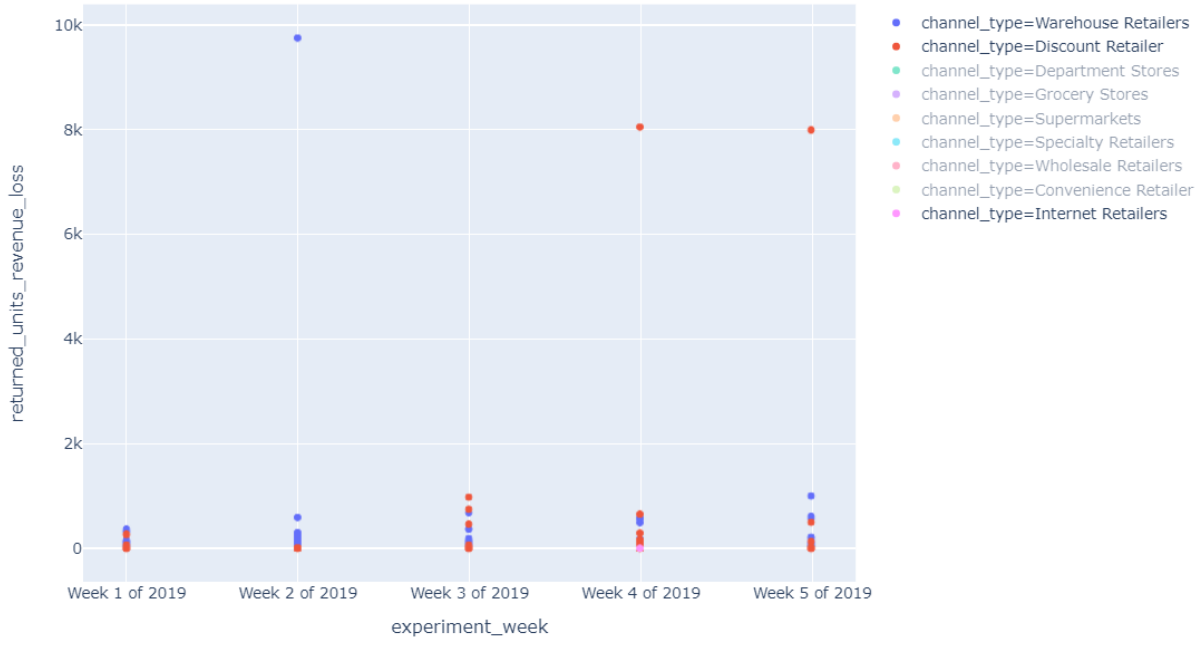
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [20]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="returned_units_revenue_loss", color="channel_type", size_max= 50)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

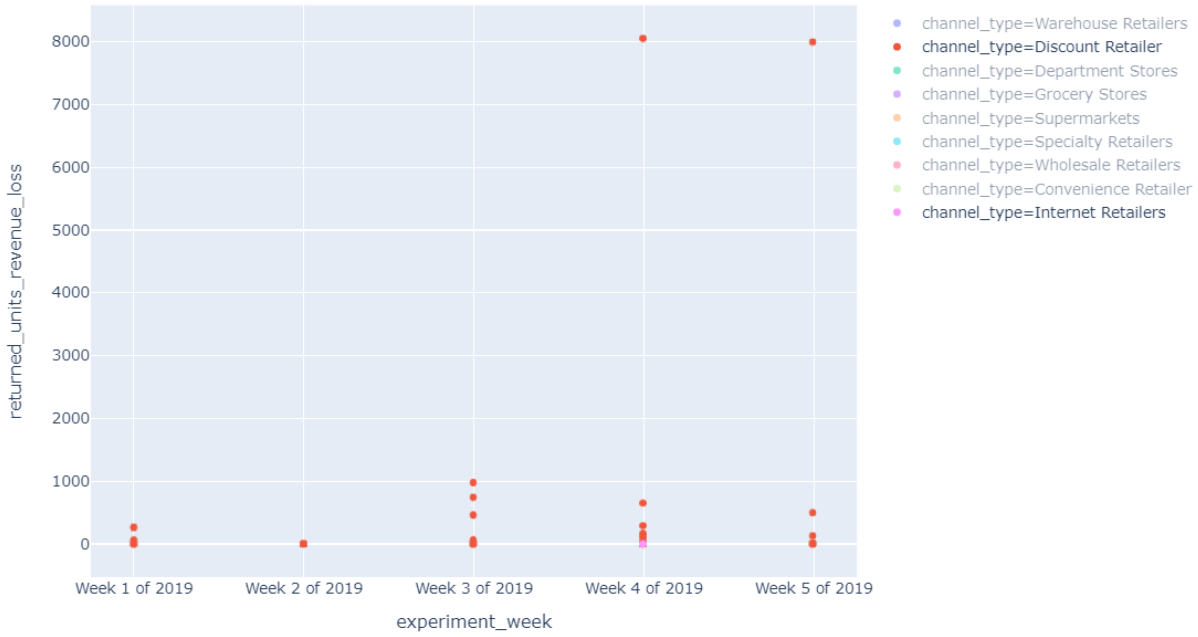
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [20]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="returned_units_revenue_loss", color="channel_type", size_max= 50)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

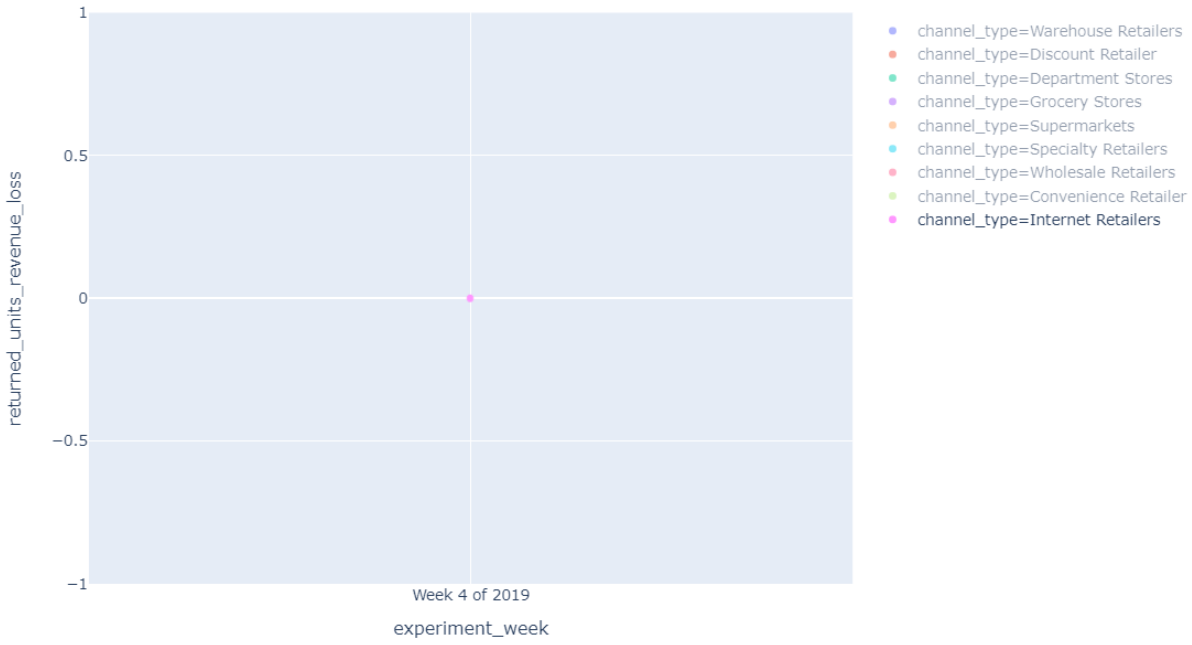
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [20]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="returned_units_revenue_loss", color="channel_type", size_max= 50)
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [20]:
```
import plotly.express as px
px.scatter(train, x="experiment_week", y="returned_units_revenue_loss", color="channel_type", size_max= 50)
```

jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

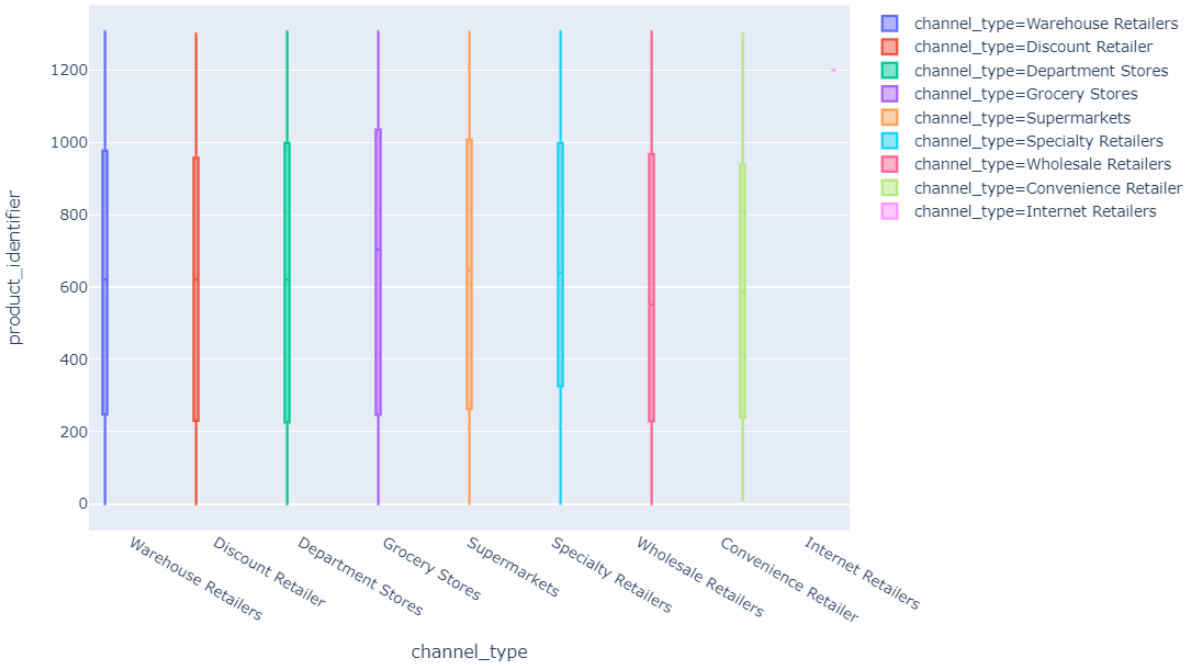
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [27]:
```
import plotly.express as px
px.box(train, x="channel_type", y="product_identifier",color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

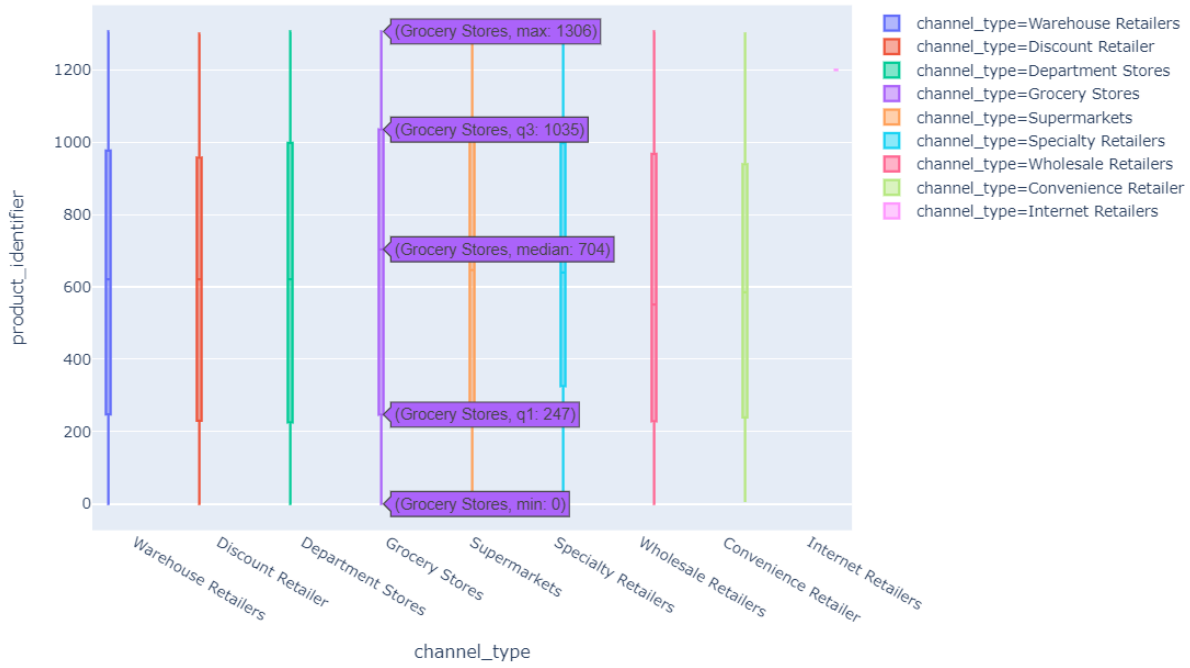
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [27]:
```
import plotly.express as px
px.box(train, x="channel_type", y="product_identifier",color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

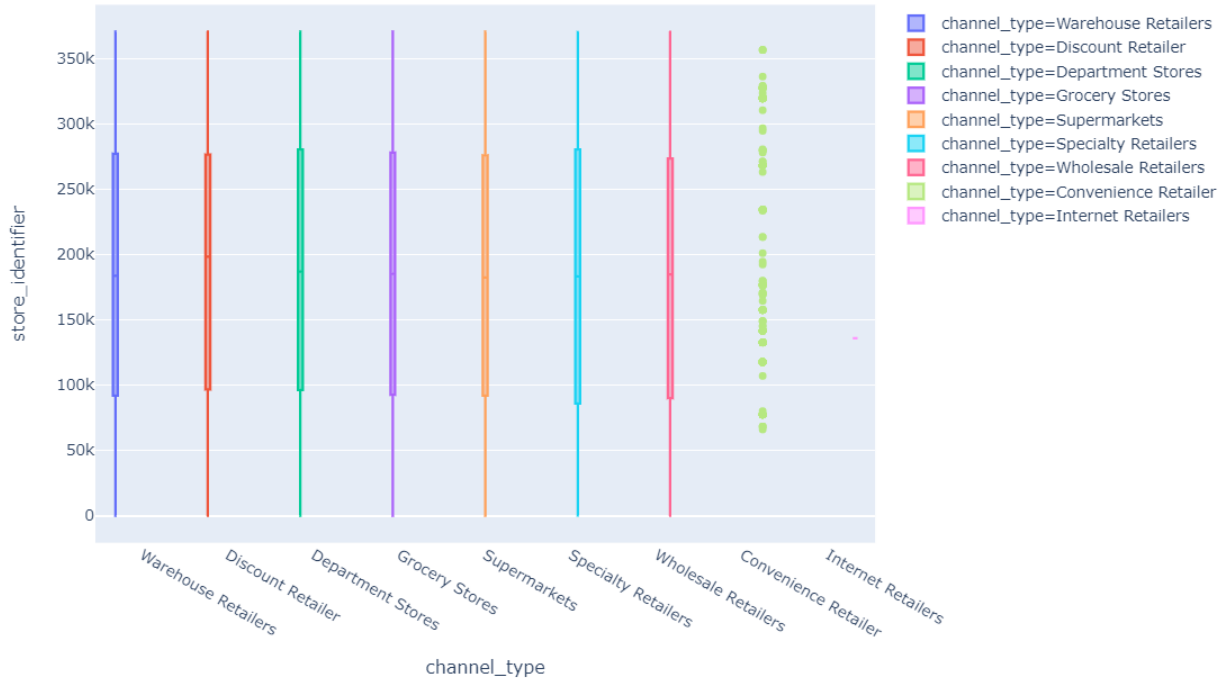
Logout

File Edit View Insert Cell Kernel Widgets Help

Notebook saved Not Trusted Python 3

Run Code

In [28]:
```
import plotly.express as px
px.box(train, x="channel_type", y="store_identifier",color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

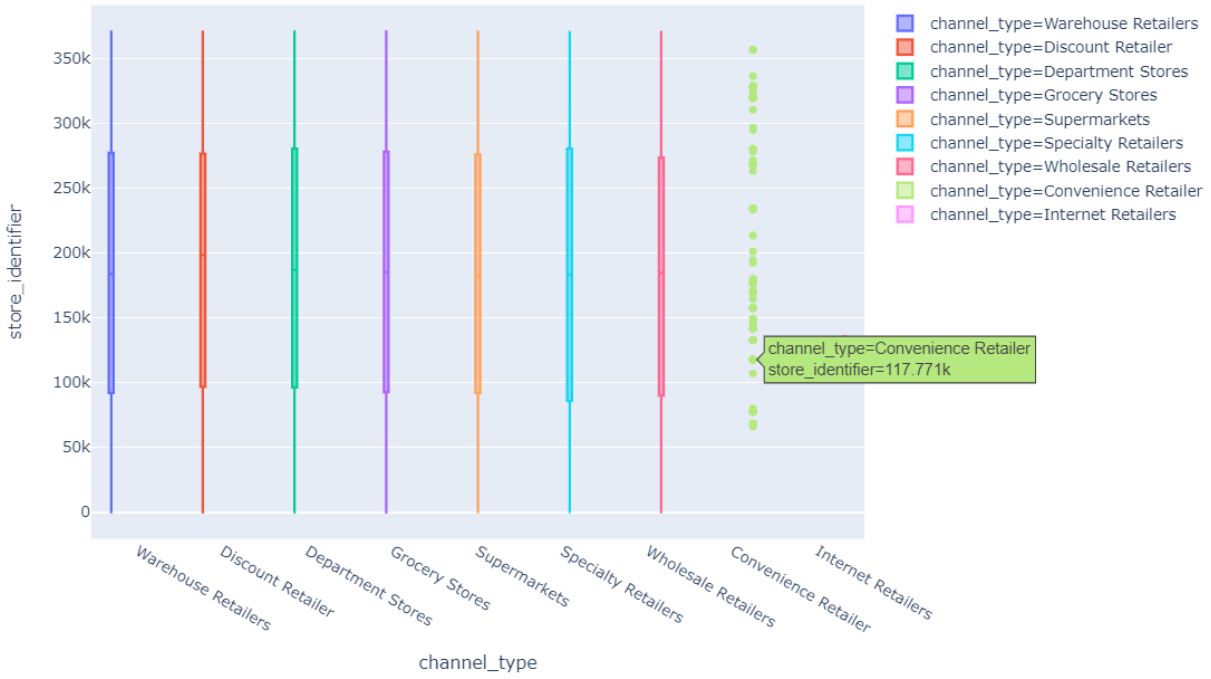
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [28]:
```
import plotly.express as px
px.box(train, x="channel_type", y="store_identifier",color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (autosaved)

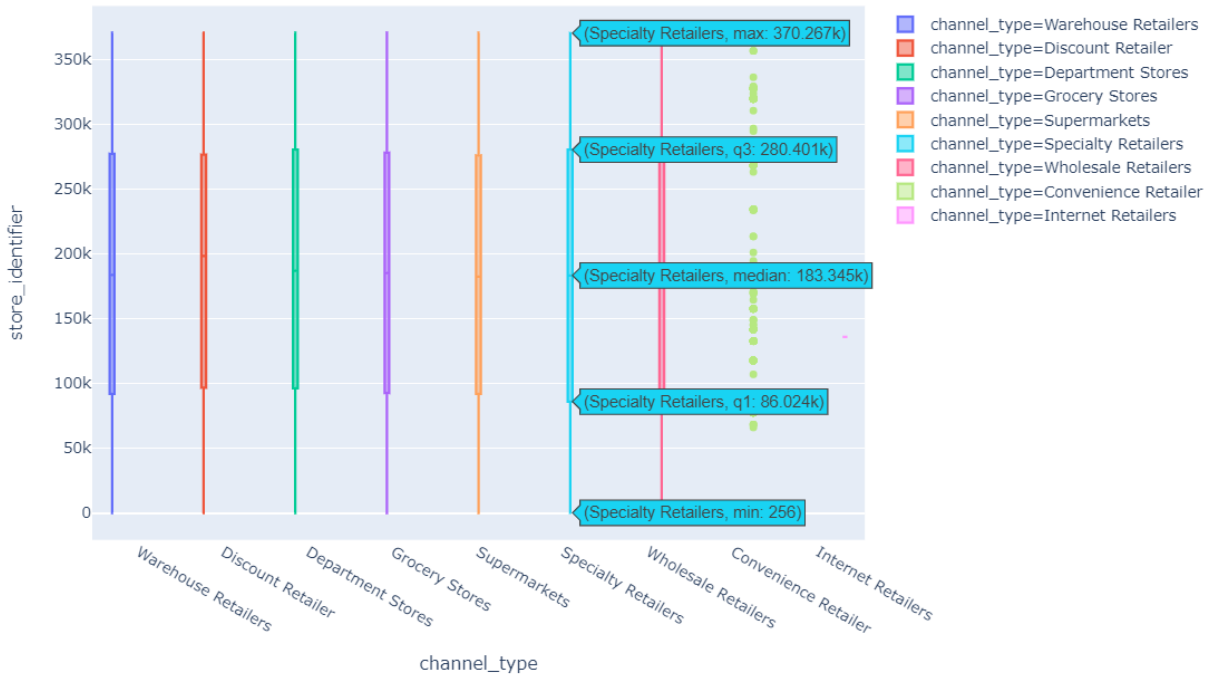
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [28]:
```
import plotly.express as px
px.box(train, x="channel_type", y="store_identifier",color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

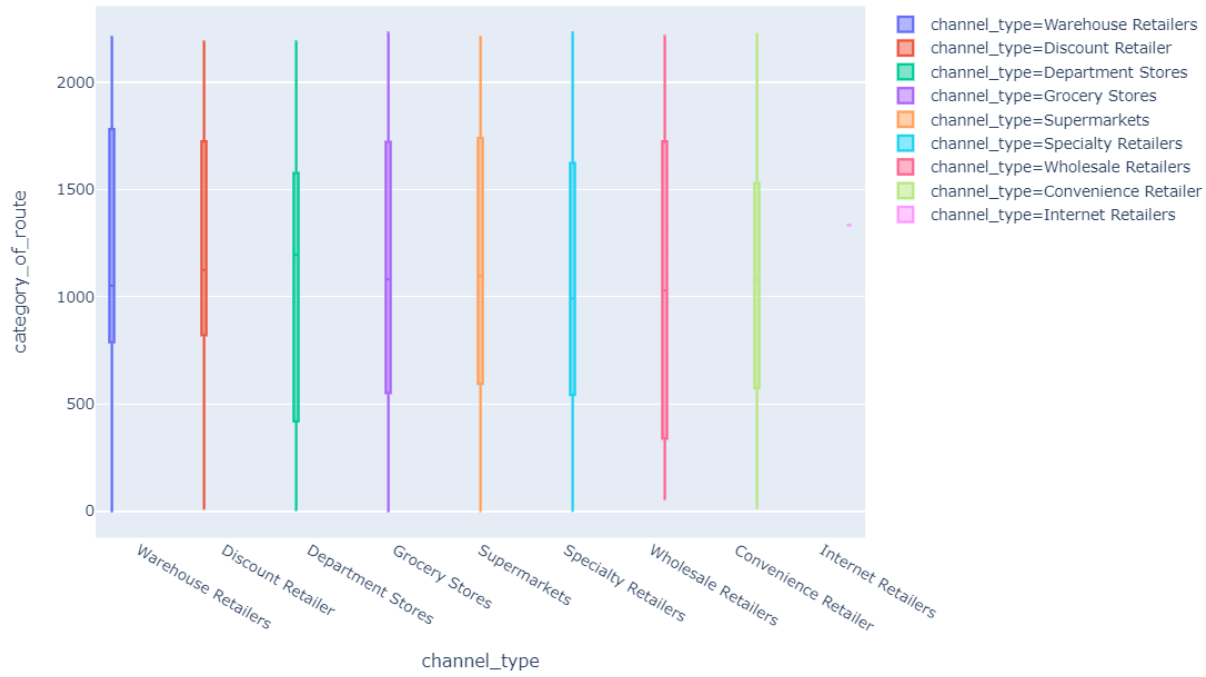
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [29]:
```
import plotly.express as px
px.box(train, x="channel_type", y="category_of_route",color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

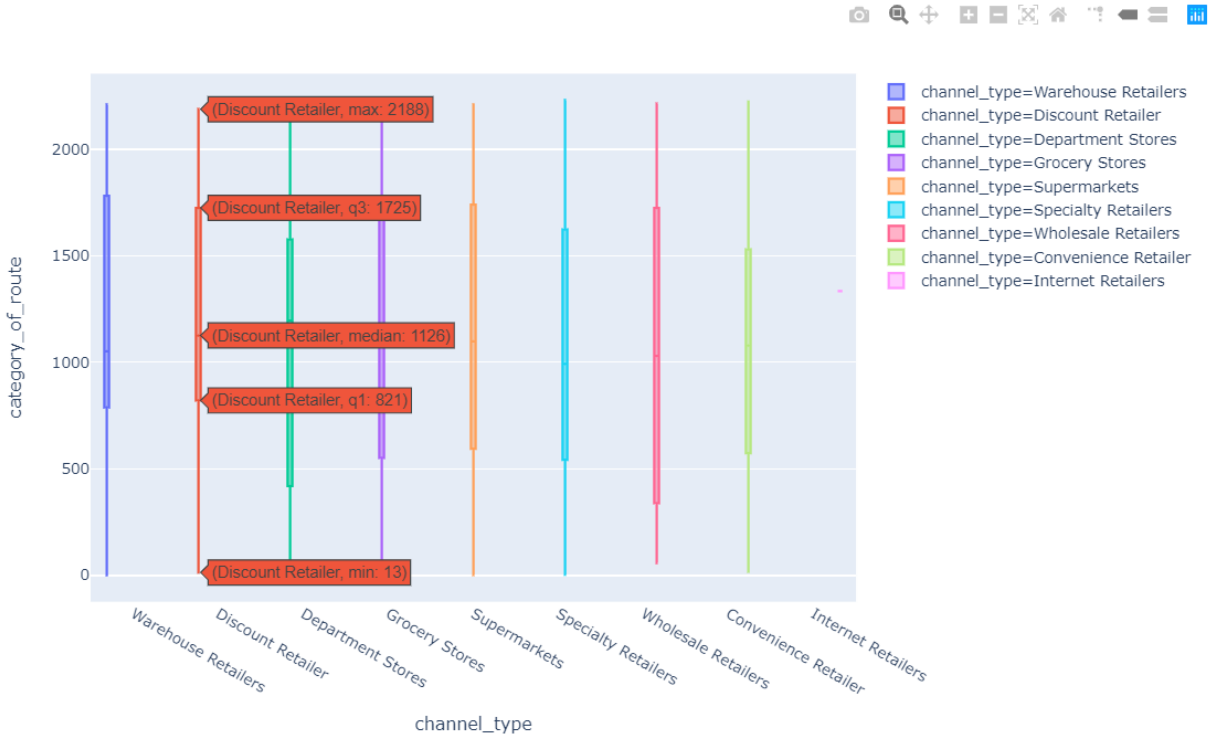
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [29]:
```
import plotly.express as px
px.box(train, x="channel_type", y="category_of_route",color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

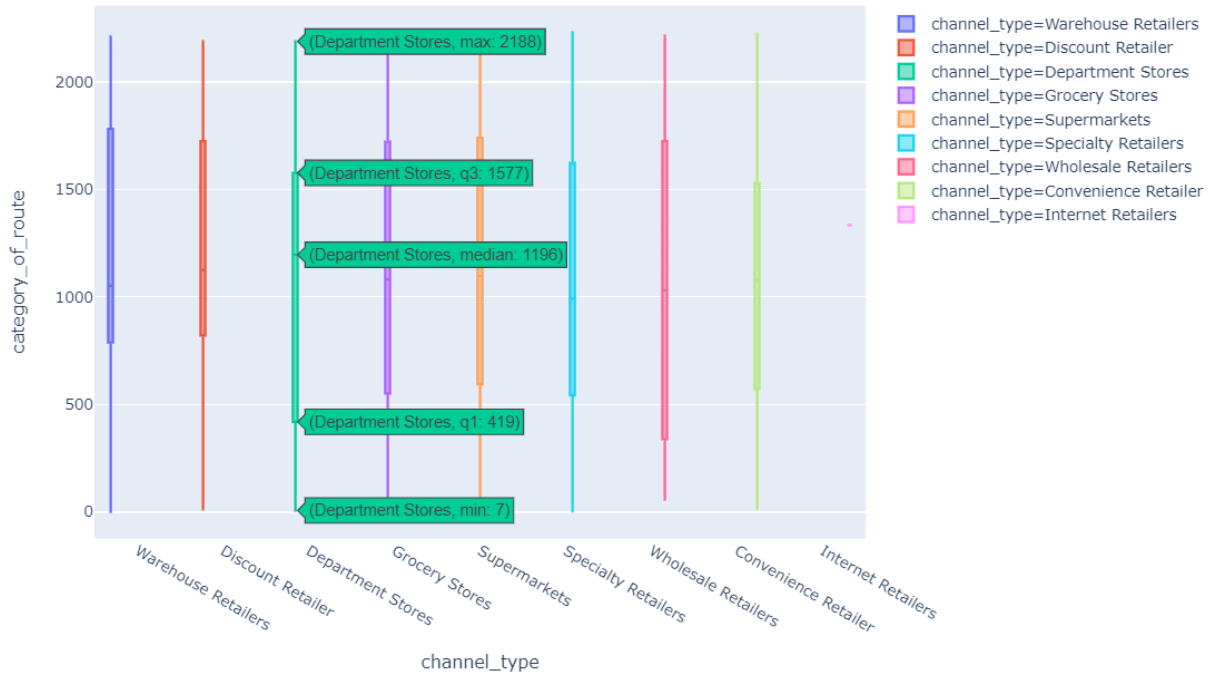
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [29]:
```
import plotly.express as px
px.box(train, x="channel_type", y="category_of_route",color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

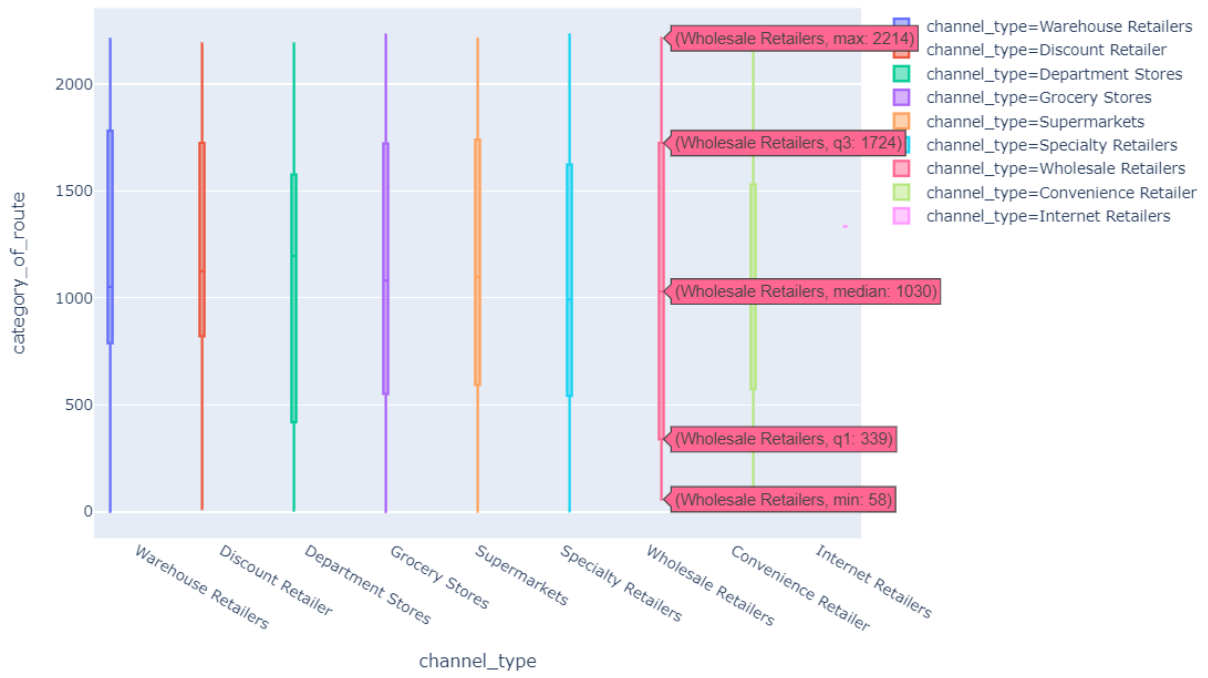
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [29]:
```
import plotly.express as px
px.box(train, x="channel_type", y="category_of_route",color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

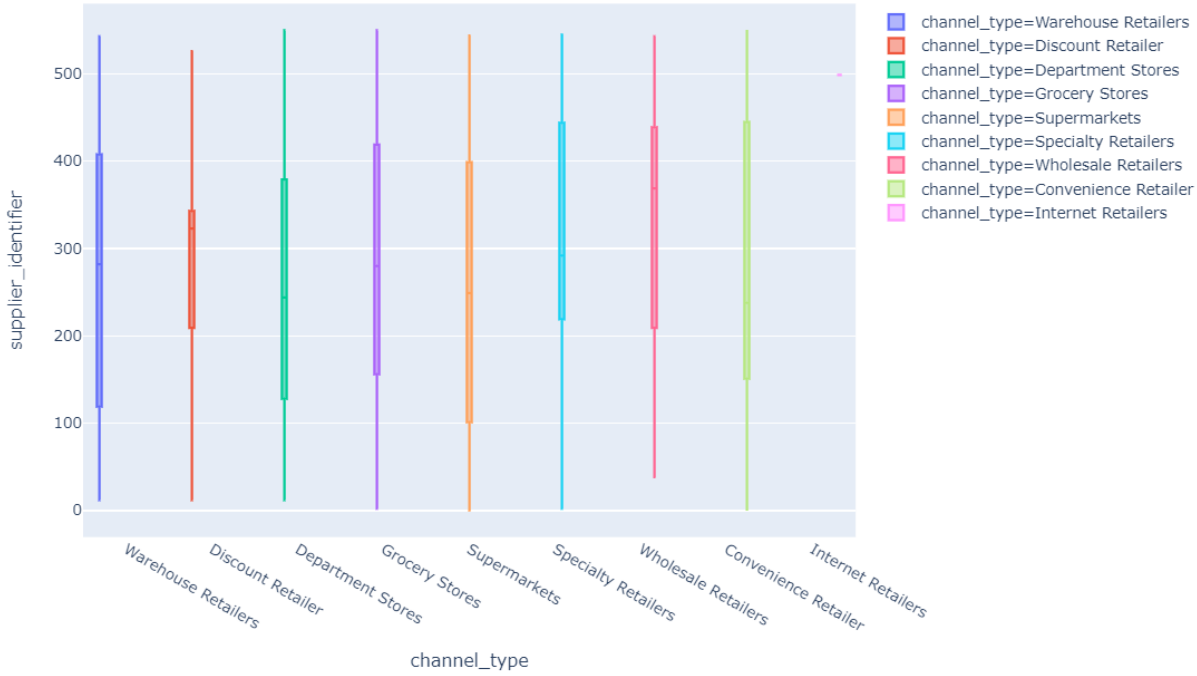
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [30]:
```
import plotly.express as px
px.box(train, x="channel_type", y="supplier_identifier",color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

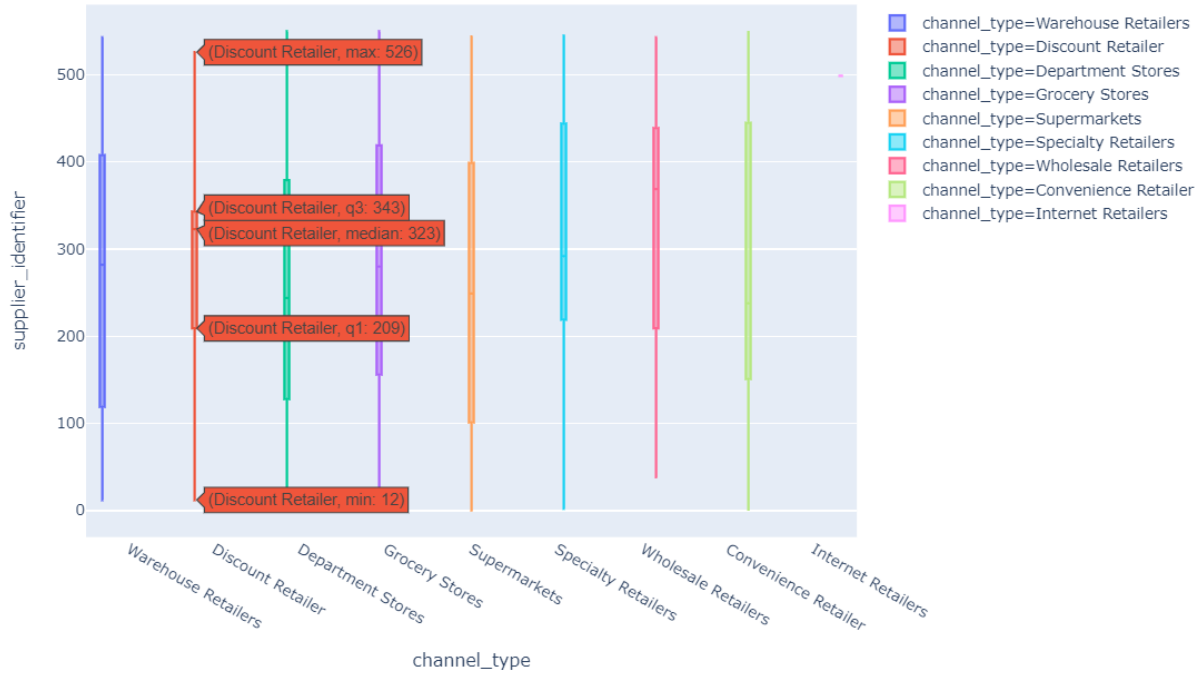
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [30]:
```
import plotly.express as px
px.box(train, x="channel_type", y="supplier_identifier",color="channel_type")
```



jupyter Shashidhar_R(Enroll_ID_3173)_visualization Last Checkpoint: 11 hours ago (unsaved changes)

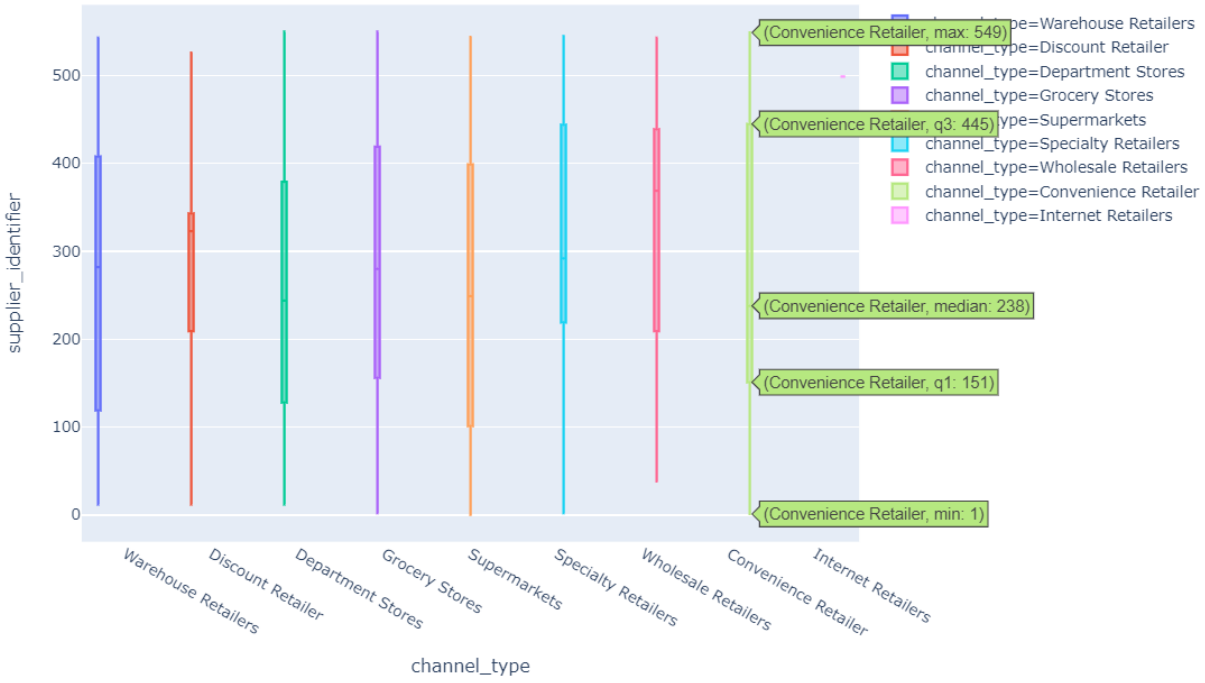
Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

Run Code

In [30]:
```
import plotly.express as px
px.box(train, x="channel_type", y="supplier_identifier",color="channel_type")
```



Data Insights:

- Based on data visualization we can clearly say that it is a regression problem.
- Here we need predict the demand of consumable retail products.
- Preprocessing and modeling is explained in detail in code part.