```python
import pandas as pd
import datetime as dt
```

```python
# import datasets

event_def = pd.read_csv('../Dataset/event_definitions.csv')
fingerhut_combined = pd.read_csv('../Dataset/fingerhut_combined.csv')
```

```
---------------------------------------------------------------------------
KeyboardInterrupt                         Traceback (most recent call last)
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/2613850120
.py in <module>
      2
      3 event_def = pd.read_csv('../Dataset/event_definitions.csv')
----> 4 fingerhut_combined = pd.read_csv('../Dataset/fingerhut_combined.csv
')

~/opt/anaconda3/lib/python3.9/site-packages/pandas/util/_decorators.py in w
rapper(*args, **kwargs)
    209                     else:
    210                         kwargs[new_arg_name] = new_arg_value
--> 211                 return func(*args, **kwargs)
    212
    213         return cast(F, wrapper)

~/opt/anaconda3/lib/python3.9/site-packages/pandas/util/_decorators.py in w
rapper(*args, **kwargs)
    329                     stacklevel=find_stack_level(),
    330                 )
--> 331             return func(*args, **kwargs)
    332
    333         # error: "Callable[[VarArg(Any), KwArg(Any)], Any]" has no

~/opt/anaconda3/lib/python3.9/site-packages/pandas/io/parsers/readers.py in
read_csv(filepath_or_buffer, sep, delimiter, header, names, index_col, usec
ols, squeeze, prefix, mangle_dupe_cols, dtype, engine, converters, true_val
ues, false_values, skipinitialspace, skiprows, skipfooter, nrows, na_values
, keep_default_na, na_filter, verbose, skip_blank_lines, parse_dates, infer
_datetime_format, keep_date_col, date_parser, dayfirst, cache_dates, iterat
or, chunksize, compression, thousands, decimal, lineterminator, quotechar,
quoting, doublequote, escapechar, comment, encoding, encoding_errors, diale
ct, error_bad_lines, warn_bad_lines, on_bad_lines, delim_whitespace, low_me
mory, memory_map, float_precision, storage_options)
    948     kwds.update(kwds_defaults)
    949
--> 950     return _read(filepath_or_buffer, kwds)
    951
    952

~/opt/anaconda3/lib/python3.9/site-packages/pandas/io/parsers/readers.py in
_read(filepath_or_buffer, kwds)
    609
    610         with parser:
```

```
--> 611            return parser.read(nrows)
    612
    613
```

~/opt/anaconda3/lib/python3.9/site-packages/pandas/io/parsers/readers.py in read(self, nrows)

```
    1776                columns,
    1777                col_dict,
->  1778            ) = self._engine.read(  # type: ignore[attr-defined
]
    1779                nrows
    1780            )
```

~/opt/anaconda3/lib/python3.9/site-packages/pandas/io/parsers/c_parser_wrapper.py in read(self, nrows)

```
    228        try:
    229            if self.low_memory:
--> 230                chunks = self._reader.read_low_memory(nrows)
    231                # destructive to chunks
    232                data = _concatenate_chunks(chunks)
```

~/opt/anaconda3/lib/python3.9/site-packages/pandas/_libs/parsers.pyx in pandas._libs.parsers.TextReader.read_low_memory()

~/opt/anaconda3/lib/python3.9/site-packages/pandas/_libs/parsers.pyx in pandas._libs.parsers.TextReader._read_rows()

~/opt/anaconda3/lib/python3.9/site-packages/pandas/_libs/parsers.pyx in pandas._libs.parsers.TextReader._convert_column_data()

~/opt/anaconda3/lib/python3.9/site-packages/pandas/_libs/parsers.pyx in pandas._libs.parsers.TextReader._convert_tokens()

~/opt/anaconda3/lib/python3.9/site-packages/pandas/_libs/parsers.pyx in pandas._libs.parsers.TextReader._convert_with_dtype()

~/opt/anaconda3/lib/python3.9/site-packages/pandas/core/dtypes/common.py in is_extension_array_dtype(arr_or_dtype)

```
    1431
    1432
->  1433 def is_extension_array_dtype(arr_or_dtype) -> bool:
    1434     """
    1435     Check if an object is a pandas extension array type.
```

KeyboardInterrupt:

In [ ]:
```
event_def
```

Out[ ]:

	event_name	journey_id	event_definition_id	milestone_number
0	application_phone_approved	1	15	1.0
1	application_phone_declined	1	16	NaN
2	application_phone_pending	1	17	NaN

3	application_web_approved	1	12	1.0	
4	application_web_declined	1	13	NaN	
5	application_web_pending	1	14	NaN	
6	application_web_submit	1	3	NaN	
7	application_web_view	1	19	NaN	
8	account_activitation	1	29	3.0	
9	account_fraud_review	1	37	NaN	
10	campaign_click	1	2	NaN	
11	customer_requested_catalog_digital	1	9	NaN	
12	fingerhut_university	1	10	NaN	
13	pre_application_3rd_party_affiliates	1	22	NaN	
14	site_registration	1	23	NaN	
15	account_downpaymentcleared	1	27	5.0	Do
16	account_downpaymentreceived	1	26	NaN	Do
17	place_downpayment	1	8	4.0	Do
18	place_downpayment_phone	1	25	4.0	Do
19	add_to_cart	1	11	NaN	
20	begin_checkout	1	6	NaN	
21	browse_products	1	4	NaN	
22	place_order_phone	1	18	2.0	
23	place_order_web	1	7	2.0	
24	view_cart	1	5	NaN	
25	order_shipped	1	28	6.0	
26	catalog_email_experian	1	20	NaN	
27	catalog_mail	1	21	NaN	

In [ ]:
```python
event_dict = {
    'Apply for Credit': [15, 17, 12, 14, 3, 19],
    'Account Activation': [29],
    'Fraud Review': [37],
    'Promotion and Discover': [1, 2, 9, 10, 22, 23, 20, 21, 24],
    'Downpayment': [27, 26, 8, 25],
    'Shopping': [11, 6, 4, 5],
    'Place Order': [18, 7],
    'Order Shipped': [28]
}
```

In [ ]:
```python
# merge event_def (event_definition_id column) with fingerhut_combined (ed

fingerhut_combined = pd.merge(fingerhut_combined, event_def, how='left', le
```

```
---------------------------------------------------------------------------
KeyboardInterrupt                         Traceback (most recent call last)
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/3321130765
.py in <module>
      1 # merge event_def (event_definition_id column) with fingerhut_combi
ned (ed_id column column)
      2
----> 3 fingerhut_combined = pd.merge(fingerhut_combined, event_def, how='l
eft', left_on='ed_id', right_on='event_definition_id')

~/opt/anaconda3/lib/python3.9/site-packages/pandas/core/reshape/merge.py in
merge(left, right, how, on, left_on, right_on, left_index, right_index, sor
t, suffixes, copy, indicator, validate)
    122         validate=validate,
    123     )
--> 124     return op.get_result(copy=copy)
    125
    126

~/opt/anaconda3/lib/python3.9/site-packages/pandas/core/reshape/merge.py in
get_result(self, copy)
    773         join_index, left_indexer, right_indexer = self._get_join_in
fo()
    774
--> 775         result = self._reindex_and_concat(
    776             join_index, left_indexer, right_indexer, copy=copy
    777         )

~/opt/anaconda3/lib/python3.9/site-packages/pandas/core/reshape/merge.py in
_reindex_and_concat(self, join_index, left_indexer, right_indexer, copy)
    735             #  necessary, but makes the `.take` more performant if
we have e.g.
    736             #  a MultiIndex for left.index.
--> 737             lmgr = left._mgr.reindex_indexer(
    738                 join_index,
    739                 left_indexer,

~/opt/anaconda3/lib/python3.9/site-packages/pandas/core/internals/managers.
```

```
py in reindex_indexer(self, new_axis, indexer, axis, fill_value, allow_dups
, copy, only_slice, use_na_proxy)
    749                 parent = None if com.all_none(*new_refs) else self
    750             else:
--> 751                 new_blocks = [
    752                     blk.take_nd(
    753                         indexer,
```

```
~/opt/anaconda3/lib/python3.9/site-packages/pandas/core/internals/managers.
py in <listcomp>(.0)
    750             else:
    751                 new_blocks = [
--> 752                     blk.take_nd(
    753                         indexer,
    754                         axis=1,
```

```
~/opt/anaconda3/lib/python3.9/site-packages/pandas/core/internals/blocks.py
in take_nd(self, indexer, axis, new_mgr_locs, fill_value)
    878
    879         # Note: algos.take_nd has upcast logic similar to coerce_to
_target_dtype
--> 880         new_values = algos.take_nd(
    881             values, indexer, axis=axis, allow_fill=allow_fill,
fill_value=fill_value
    882         )
```

```
~/opt/anaconda3/lib/python3.9/site-packages/pandas/core/array_algos/take.py
in take_nd(arr, indexer, axis, fill_value, allow_fill)
    115
    116         arr = np.asarray(arr)
--> 117         return _take_nd_ndarray(arr, indexer, axis, fill_value,
allow_fill)
    118
    119
```

```
~/opt/anaconda3/lib/python3.9/site-packages/pandas/core/array_algos/take.py
in _take_nd_ndarray(arr, indexer, axis, fill_value, allow_fill)
    161             arr.ndim, arr.dtype, out.dtype, axis=axis, mask_info=mask_i
nfo
    162         )
--> 163     func(arr, indexer, out, fill_value)
    164
    165     if flip_order:
```

```
KeyboardInterrupt:
```

```python
In [ ]:    # keep only rows where combined_id is between 1 and 1000 (first 1000 custom

           fingerhut_small = fingerhut_combined[(fingerhut_combined['combined_id'] >=
```

```python
In [ ]:    # groupby combined_id and only keep groups that has at least one instance

           fingerhut_active_account = fingerhut_small.groupby('combined_id').filter(la
```

```python
# only keep 'combined_id', 'event_timestamp', 'event_name_y'

fingerhut_active_account = fingerhut_active_account[['combined_id', 'ed_id'
```

```python
fingerhut_active_account
```

	combined_id	ed_id
31	1	2
32	1	4
33	1	11
34	1	4
35	1	4
...	...	...
31905	999	1
31906	999	1
31907	999	5
31908	999	5
31909	999	28

11930 rows × 2 columns

```python
# Use the event_dict to create a new column called 'event_type' in fingerhu

fingerhut_active_account['event_type'] = fingerhut_active_account['ed_id'].
```

```python
# remove ed_id column

fingerhut_active_account = fingerhut_active_account.drop(columns=['ed_id'])

# remove duplicates

fingerhut_active_account = fingerhut_active_account.drop_duplicates()
```

```python
fingerhut_active_account
```

Out[ ]:

	combined_id	event_type
31	1	Promotion and Discover
32	1	Shopping
57	1	Apply for Credit
82	1	Account Activation
92	1	Downpayment
...	...	...
31893	999	Place Order
31894	999	Downpayment
31896	999	Account Activation
31905	999	Promotion and Discover
31909	999	Order Shipped

1467 rows × 2 columns

In [ ]:

```python
# groupby combined_id, and if the last event_type is not 'Order Shipped',

fingerhut_active_account = fingerhut_active_account.groupby('combined_id').

# reset index

fingerhut_active_account = fingerhut_active_account.reset_index(drop=True)
```

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:
```

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391 4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391 4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391 4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391 4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391 4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391 4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391 4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391 4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391 4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391 4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a future version. Use pandas.concat instead.

```
/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.

/var/folders/6m/88dwrhnx7m3cybxwl1p0rtq40000gn/T/ipykernel_74297/102283391
4.py:3: FutureWarning:

The frame.append method is deprecated and will be removed from pandas in a
future version. Use pandas.concat instead.
```

In [ ]:
```python
# check for NA values

fingerhut_active_account.isna().sum()
```

Out[ ]:
```
combined_id     0
event_type      0
dtype: int64
```

In [ ]:
```python
fingerhut_active_account['sequence'] = fingerhut_active_account.groupby('co
```

In [ ]:
```python
fingerhut_active_account
```

Out[ ]:

	combined_id	event_type	sequence
0	1	Promotion and Discover	1
1	1	Shopping	2
2	1	Apply for Credit	3
3	1	Account Activation	4
4	1	Downpayment	5
...	...	...	...
1509	999	Place Order	3
1510	999	Downpayment	4
1511	999	Account Activation	5
1512	999	Promotion and Discover	6
1513	999	Order Shipped	7

1514 rows × 3 columns

In [ ]:
```python
pivot_fingerhut_active_account = fingerhut_active_account.pivot(index='comk
```

In [ ]:
```python
pivot_fingerhut_active_account = pivot_fingerhut_active_account.fillna('')
```

In [ ]:
```python
pivot_fingerhut_active_account[50:100]
```

Out[ ]:

sequence	combined_id	1	2	3	4	5
50	211	Promotion and Discover	Apply for Credit	Shopping	Downpayment	Place Order
51	212	Promotion and Discover	Apply for Credit	Shopping	Downpayment	Place Order
52	216	Apply for Credit	Shopping	Promotion and Discover	Account Activation	Downpayment
53	221	Apply for Credit	Shopping	Promotion and Discover	Account Activation	No Order Made
54	224	Shopping	Account Activation	Downpayment	Place Order	Apply for Credit
55	225	Promotion and Discover	Apply for Credit	Account Activation	Place Order	Downpayment
56	233	Apply for Credit	Shopping	Promotion and Discover	Account Activation	No Order Made

57	234	Apply for Credit	Shopping	Promotion and Discover	Account Activation	No Order Made	
58	235	Apply for Credit	Shopping	Promotion and Discover	Account Activation	Place Order	D
59	236	Account Activation	Apply for Credit	Shopping	Place Order	Downpayment	
60	238	Promotion and Discover	Apply for Credit	Shopping	Account Activation	Place Order	D
61	239	Account Activation	Promotion and Discover	Apply for Credit	Shopping	No Order Made	
62	240	Apply for Credit	Shopping	Promotion and Discover	Place Order	Downpayment	
63	242	Promotion and Discover	Apply for Credit	Shopping	Order Shipped	Account Activation	
64	244	Promotion and Discover	Apply for Credit	Shopping	Downpayment	Place Order	
65	246	Apply for Credit	Shopping	Downpayment	Place Order	Promotion and Discover	
66	248	Apply for Credit	Shopping	Promotion and Discover	Place Order	Downpayment	
67	249	Promotion and Discover	Apply for Credit	Shopping	Account Activation	Downpayment	
68	252	Promotion and Discover	Shopping	Apply for Credit	Account Activation	Place Order	D
69	257	Apply for Credit	Promotion and Discover	Shopping	Account Activation	Downpayment	
70	258	Promotion and Discover	Apply for Credit	Place Order	Shopping	Account Activation	D
71	259	Apply for Credit	Shopping	Promotion and Discover	Account Activation	Downpayment	
72	264	Apply for Credit	Shopping	Promotion and Discover	Downpayment	Place Order	
73	265	Account Activation	Promotion and Discover	Apply for Credit	Shopping	Place Order	
74	281	Promotion and Discover	Apply for Credit	Shopping	Account Activation	Downpayment	

75	283	Promotion and Discover	Apply for Credit	Shopping	Account Activation	Downpayment
76	295	Promotion and Discover	Apply for Credit	Shopping	Downpayment	Place Order
77	297	Promotion and Discover	Apply for Credit	Shopping	Account Activation	No Order Made
78	300	Apply for Credit	Shopping	Promotion and Discover	Place Order	Downpayment
79	304	Apply for Credit	Shopping	Account Activation	Promotion and Discover	Place Order D
80	307	Apply for Credit	Place Order	Promotion and Discover	Account Activation	Downpayment
81	321	Promotion and Discover	Apply for Credit	Shopping	Account Activation	No Order Made
82	326	Account Activation	Promotion and Discover	Apply for Credit	Shopping	No Order Made
83	328	Shopping	Apply for Credit	Promotion and Discover	Downpayment	Place Order
84	332	Apply for Credit	Promotion and Discover	Shopping	Account Activation	Place Order D
85	339	Promotion and Discover	Apply for Credit	Shopping	Account Activation	Place Order D
86	344	Apply for Credit	Promotion and Discover	Shopping	Account Activation	No Order Made
87	345	Account Activation	Apply for Credit	Shopping	Promotion and Discover	Downpayment
88	347	Promotion and Discover	Apply for Credit	Shopping	Account Activation	Place Order D
89	350	Account Activation	Apply for Credit	Promotion and Discover	Shopping	Place Order D
90	361	Account Activation	Promotion and Discover	Apply for Credit	Shopping	Place Order D
91	362	Promotion and Discover	Shopping	Apply for Credit	Downpayment	Place Order
92	365	Apply for Credit	Shopping	Account Activation	Downpayment	Place Order

93	369	Account Activation	Apply for Credit	Shopping	Promotion and Discover	Downpayment
94	377	Promotion and Discover	Apply for Credit	Shopping	Account Activation	Place Order D
95	378	Promotion and Discover	Apply for Credit	Shopping	Place Order	Downpayment
96	388	Apply for Credit	Promotion and Discover	Shopping	Account Activation	Downpayment
97	389	Apply for Credit	Shopping	Account Activation	Downpayment	Place Order
98	390	Apply for Credit	Shopping	Place Order	Downpayment	Account Activation
99	395	Promotion and Discover	Apply for Credit	Account Activation	Shopping	Downpayment

In [ ]:
```python
len(pivot_fingerhut_active_account)
```

Out[ ]:
```
230
```

In [ ]:
```python
import pandas as pd
import plotly.graph_objects as go

def generate_sankey_chart_data(df: pd.DataFrame):
    # list of list: each list is the set of nodes in each tier/column
    column_values = [df[col] for col in df.columns]

    # this generates the labels for the sankey by taking all the unique va
    labels = sum([list(node_values.unique()) for node_values in column_valu

    # initializes a dict of dicts (one dict per tier)
    link_mappings = {col: {} for col in df.columns}

    # each dict maps a node to a unique number value
    i = 0
    for col, nodes in zip(df.columns, column_values):
        for node in nodes.unique():
            link_mappings[col][node] = i
            i += 1

    # specifying which columns are serving as sources and which as targets
    source_cols = df.columns[:-1]
    target_cols = df.columns[1:]
    links = []
```

```python
        # loop to create a list of links in the format [((src, tgt), wt), (),
        for source_col, target_col in zip(source_cols, target_cols):
            for source, target in zip(df[source_col], df[target_col]):
                links.append(
                    (
                        link_mappings[source_col][source],
                        link_mappings[target_col][target],
                        1  # Weight is 1 for counting transitions
                    )
                )

        # creating a dataframe with 3 columns: source, target, and weight
        df_links = pd.DataFrame(links, columns=["source", "target", "weight"])

        # generating three lists needed for the sankey visual
        sources = df_links["source"]
        targets = df_links["target"]
        weights = df_links["weight"]

        return labels, sources, targets, weights


# Your DataFrame
df = pivot_fingerhut_active_account.iloc[:, 2:]

# Call the generate_sankey_chart_data function
labels, sources, targets, weights = generate_sankey_chart_data(df=df)


# Map colors to labels
label_colors = {
    'Promotion and Discover': 'pink',
    'Apply for Credit': 'green',
    'Shopping': 'blue',
    'Account Activation': 'cyan',
    'Downpayment': 'purple',
    'Place Order': 'yellow',
    'Order Shipped': 'orange',
    'No Order Made': 'red',
    '': 'black'  # Adjust this for labels with empty strings
}

# Create the Sankey diagram
fig = go.Figure(data=[go.Sankey(
    node=dict(
        pad=15,
        thickness=20,
        line=dict(color="black", width=0.5),
        label=labels,
        color=[label_colors[label] for label in labels]  # Map colors based
    ),
    link=dict(
        source=sources,
        target=targets,
        value=weights
    )
)])
```

```
fig.update_layout(title_text="Journey flow of customers with account activa
fig.show()
```

Key points:

- A large proportion of customers exposed to the promotion and discover stage move onto: Apply for Credit, Account Activation, Downpayment
- A large proportion of customers move onto Downpayment or Place Order after Account Activation, which may indicate that Account Activation is a significant factor to the customer's decision to purchase. This will require further investigation and hypothesis testing. If it holds to be true then we can focus more on increaseing Account Activation rate than other stages.