

# U.S. Meat Production Analysis – anomalies and Transformers

**Target Audience:**  
Highly Technical, but  
unaware of my project

**Caleb Hart**

STAT443 Statistical Machine Learning II – Prof. Joseph Reid

# Intro

- This report examines U.S. meat production statistics and attempts to forecast meat production for the year.
- The sanitized dataset includes variables such as beef, veal, pork, lamb and mutton, broilers, turkey, total poultry production, and total meat production.
- Dataset for the transformer forecasting looks a little different this time

# Source

- [USDA](#)
- Meat statistics table, historical
- Cropped the data to the period of January 2025 – January 1977 as this was the period that lacked any missing values for federally inspected meat.
- I additional removed all columns for non-federally inspected meat as I wanted to look at the longest historical trends.
- Variables casted to float32 and commas removed (excel moment).

# Quick Look(uncleaned)

Red meat and poultry production (million pounds)															
Type 1/	Commercial 2/					Federally inspected									
	Beef 3/	Veal 3/	Pork 3/	Lamb and mutton 3/	Total red meat 3/ 4/	Beef 3/	Veal 3/	Pork 3/	Lamb and mutton 3/	Total red meat 3/ 4/	Broilers 5/	Other chicken 5/	Turkey 5/	Total poultry 4/ 5/ 6/	Total red meat and poultry 4/
Jan 2025	2,370.1	2.5	2,501.9	10.9	4,885.3	2,335.6	2.3	2,492.6	9.7	4,840.2	4,137.7	48.0	415.2	4,612.7	9,452.9
Jan 2024	2,280.8	3.9	2,472.7	10.6	4,768.0	2,246.1	3.8	2,462.4	9.6	4,721.9	4,051.4	47.4	435.2	4,547.6	9,269.5
Jan-2025	2,370.1	2.5	2,501.9	10.9	4,885.3	2,335.6	2.3	2,492.6	9.7	4,840.2	4,137.7	48.0	415.2	4,612.7	9,452.9
Dec-2024	2,200.7	3.1	2,328.7	12.0	4,544.4	2,168.6	2.9	2,319.9	10.8	4,502.1	3,882.4	42.0	376.3	4,311.2	8,813.3
Nov-2024	2,216.6	3.1	2,334.4	10.4	4,564.5	2,185.0	3.0	2,325.7	9.2	4,522.9	3,651.4	36.9	394.1	4,093.1	8,616.0
Oct-2024	2,465.5	3.3	2,543.5	11.7	5,024.0	2,425.8	3.1	2,532.2	10.3	4,971.4	4,370.9	52.6	493.9	4,929.6	9,901.0
Sep-2024	2,204.9	3.1	2,232.6	10.6	4,451.2	2,170.7	2.9	2,221.8	9.4	4,404.8	3,892.6	49.2	413.4	4,366.2	8,771.0
Aug-2024	2,288.1	3.1	2,289.9	10.5	4,591.7	2,256.1	3.0	2,278.5	9.2	4,546.8	4,038.7	49.6	435.0	4,533.2	9,080.0
Jul-2024	2,286.8	3.3	2,252.8	11.3	4,554.3	2,255.9	3.2	2,242.5	10.0	4,511.7	4,072.1	50.5	433.5	4,566.2	9,077.9
Jun-2024	2,135.8	3.1	2,117.7	10.4	4,267.0	2,103.6	3.0	2,109.2	9.1	4,225.0	3,724.9	47.7	408.9	4,190.7	8,415.7
May-2024	2,326.4	3.5	2,278.6	11.5	4,619.9	2,292.3	3.4	2,269.5	10.2	4,575.4	4,011.0	50.3	446.8	4,518.9	9,094.3
Apr-2024	2,303.4	3.5	2,317.4	11.4	4,635.6	2,269.3	3.3	2,308.2	10.1	4,590.9	3,917.7	47.2	450.7	4,427.5	9,018.4
Mar-2024	2,110.4	3.6	2,250.1	12.3	4,376.4	2,076.5	3.4	2,240.7	11.1	4,331.7	3,636.8	42.9	409.5	4,100.3	8,432.0
Feb-2024	2,168.6	3.6	2,371.2	10.9	4,554.3	2,135.0	3.4	2,361.4	9.8	4,509.7	3,741.9	45.0	423.9	4,221.4	8,731.1
Jan-2024	2,280.8	3.9	2,472.7	10.6	4,768.0	2,246.1	3.8	2,462.4	9.6	4,721.9	4,051.4	47.4	435.2	4,547.6	9,269.5

# Yesterday's Data

```
[7]: print(df.sample(n=10))
```

	Date	Beef	Veal	Pork	Lamb and mutton	Total red meat	Broilers	\
4	24-Sep	2204.9	3.1	2232.6	10.6	4451.2	3892.6	
268	2-Sep	2201.0	16.3	1638.0	17.6	3873.2	2596.4	
217	6-Dec	2049.6	13.1	1796.4	15.4	3874.5	2737.8	
105	16-Apr	1964.0	6.0	2000.2	12.9	3983.0	3302.1	
0	25-Jan	2370.1	2.5	2501.9	10.9	4885.3	4137.7	
122	14-Nov	1849.5	6.5	1890.2	11.5	3757.7	2939.9	
277	1-Dec	2110.0	16.0	1668.0	19.0	3813.6	2464.8	
139	13-Jun	2157.0	8.4	1676.3	12.7	3854.5	3002.0	
210	7-Jul	2256.7	10.7	1659.5	13.5	3940.4	3055.0	
285	1-Apr	1939.0	15.0	1533.0	20.0	3507.5	2515.7	

	Other chicken	Turkey	Total poultry	Total red meat and poultry
4	49.2	413.4	4366.2	8771.0
268	46.4	444.6	3098.3	6918.3
217	36.9	429.1	3215.5	7043.0
105	43.8	484.8	3842.1	7787.9
0	48.0	415.2	4612.7	9452.9
122	36.1	480.8	3467.3	7190.1
277	40.9	419.8	2935.4	6694.4
139	42.9	473.4	3527.6	7345.1
210	44.6	505.2	3616.0	7512.8
285	42.2	428.8	2997.9	6455.9

# Anomaly Detection

	Z_Score_Anomaly	IsolationForest_Anomaly	DBSCAN_Anomaly
2	False	True	True
16	True	True	True
28	False	True	True
50	False	True	False
62	False	True	True
74	False	False	True
193	False	False	True
194	False	False	True
206	True	True	True
213	False	False	True
238	False	True	True
249	True	True	True
250	False	True	True
253	False	True	True
266	False	True	True
273	True	True	True
278	False	True	True
285	True	True	True
286	False	True	True

	Beef	Veal	Pork	Lamb and mutton	Total red meat	Broilers \
2	1.763860	-1.506932	2.097974	-0.794782	2.335388	2.662100
16	1.133685	-1.283683	1.172164	-1.240029	1.360655	2.154742
28	2.016554	-1.159655	1.202427	-0.794782	1.782946	2.283181
50	1.818142	-1.010823	2.303833	-1.240029	2.516370	1.425102
62	1.600389	-0.564324	2.332975	-0.349536	2.451574	2.020378
74	1.546106	-0.539519	1.750515	-0.228105	1.992975	1.567523
193	-1.394503	0.477505	-0.358148	-0.349536	-0.889050	-1.385869
194	0.986436	1.023225	0.664800	0.540958	0.953725	0.054689
206	1.623474	0.427894	0.609506	1.107636	1.194755	0.322942
213	-0.403069	1.023225	-0.451925	2.402899	-0.489940	-0.757654
238	-2.591835	0.651143	-1.317956	0.581435	-2.133859	-1.373546
249	-0.466710	1.172058	-0.676091	3.455301	-0.678184	-0.815949
250	-2.363475	1.271280	-1.535398	0.460004	-2.186924	-1.847729
253	-2.494501	1.444918	-0.986189	1.067159	-1.829709	-1.951760
266	2.053990	2.313109	-0.564008	2.443376	0.542606	-0.697226
273	-0.772439	1.395307	-1.498037	3.455301	-1.424734	-1.569760
278	1.286549	2.139471	-0.537855	2.564807	0.215833	-0.830167
285	-0.541583	1.643362	-1.329911	3.779117	-1.188452	-1.561465
286	-1.883044	1.643362	-1.852967	1.350498	-2.200610	-2.192761

	Other chicken	Turkey	Total poultry	Total red meat and poultry \
2	2.059533	0.525361	2.675792	2.601590
16	3.522404	0.417453	2.183877	1.868369
28	2.717825	-0.711434	2.206174	2.075121
50	1.181810	1.324989	1.515216	2.024943
62	0.572281	2.113550	2.169013	2.373033
74	1.645053	2.711196	1.786722	1.941008
193	-1.353833	0.929326	-1.297850	-1.145038
194	1.571909	2.957448	0.316319	0.617610
206	0.840474	3.048755	0.582714	0.879268
213	-1.134402	0.193336	-0.736027	-0.646941
238	-1.305071	-1.818185	-1.513615	-1.856323
249	0.230944	-0.111021	-0.804774	-0.780792
250	-2.889848	-2.128076	-2.018536	-2.169998
253	-2.499749	-0.537120	-1.983001	-1.981767
266	1.596290	1.410762	-0.548133	-0.076074
273	-0.329823	-0.766771	-1.606284	-1.586127
278	0.694187	1.836862	-0.650093	-0.284796
285	0.060276	-0.357273	-1.560066	-1.450963
286	-1.085640	-1.815419	-2.314196	-2.349035

# Anomalies Continued

```
Proportion of anomalies detected by Isolation Forest: 5.21%  
Proportion of anomalies detected by DBSCAN: 6.25%  
Proportion of anomalies detected by Z-Score: 1.74%
```

```
### METHOD 1: Z-Score ###  
z_scores = np.abs(stats.zscore(df_scaled))  
threshold = 3  
df_scaled['Z_Score_Anomaly'] = (z_scores > threshold).any(axis=1)  
  
### METHOD 2: Isolation Forest ###  
iso_forest = IsolationForest(contamination=0.05, random_state=42)  
df_scaled['IsolationForest_Anomaly'] = iso_forest.fit_predict(df_scaled) == -1 # Convert to True/False  
  
### METHOD 3: DBSCAN ###  
dbscan = DBSCAN(eps=1.5, min_samples=5)  
df_scaled['DBSCAN_Anomaly'] = dbscan.fit_predict(df_scaled) == -1 # Convert to True/False
```

# New Clean

```
df = pd.read_csv(r'T:\Users\caleb.hart\DSAI443\Meats1977.csv')

# Apply conversion by removing commas and turning to numeric
for col in df:
    if df[col].dtype == 'object': # Check if the column contains string values
        df[col] = df[col].replace({' ',''}, regex=True) # Remove commas
        df[col] = pd.to_numeric(df[col], errors='raise') # Convert to numeric
df.head()
```

	Beef 3/	Veal 3/	Pork 3/	Lamb and mutton 3/	Total red meat 3/ 4/	Broilers 5/	Turkey 5/	Total poultry 4/ 5/ 6/	Total red meat and poultry 4/
0	2335.6	2.3	2492.6	9.7	4840.2	4137.7	415.2	4612.7	9452.9
1	2246.1	3.8	2462.4	9.6	4721.9	4051.4	435.2	4547.6	9269.5
2	2335.6	2.3	2492.6	9.7	4840.2	4137.7	415.2	4612.7	9452.9
3	2168.6	2.9	2319.9	10.8	4502.1	3882.4	376.3	4311.2	8813.3
4	2185.0	3.0	2325.7	9.2	4522.9	3651.4	394.1	4093.1	8616.0

Units: Millions of  
pounds



# Data Description

- Variables: Beef, Veal, Pork, Lamb and Mutton, Total Red Meat, Broilers, Turkey, Total Poultry, Total Red Meat and Poultry, and Date.
- Variable Types: Continuous Numeric (e.g., production volume)

# Transformer-based supervised forecasting

```
class TransformerForecastingModel(torch.nn.Module):
    def __init__(self, input_dim, embed_dim, num_heads, ff_dim, num_layers, dropout=0.1):
        super().__init__()
        self.embedding = torch.nn.Linear(input_dim, embed_dim)
        self.relu1 = torch.nn.ReLU()
        self.dropout1 = torch.nn.Dropout(dropout)

        encoder_layer = torch.nn.TransformerEncoderLayer(
            d_model=embed_dim, nhead=num_heads, dim_feedforward=ff_dim, dropout=dropout
        )
        self.transformer_encoder = torch.nn.TransformerEncoder(encoder_layer, num_layers=num_layers)

        self.relu2 = torch.nn.ReLU()
        self.dropout2 = torch.nn.Dropout(dropout)
        self.fc_out = torch.nn.Linear(embed_dim, input_dim)
        self.relu3 = torch.nn.ReLU()

    def forward(self, x):
        x = self.embedding(x)
        x = self.relu1(x)
        x = self.dropout1(x)
        x = self.transformer_encoder(x)
        x = self.relu2(x)
        x = self.dropout2(x)
        x = self.fc_out(x[:, -1, :]) # Take the last time step output
        x = self.relu3(x)
        return x
```

# Model Parameters

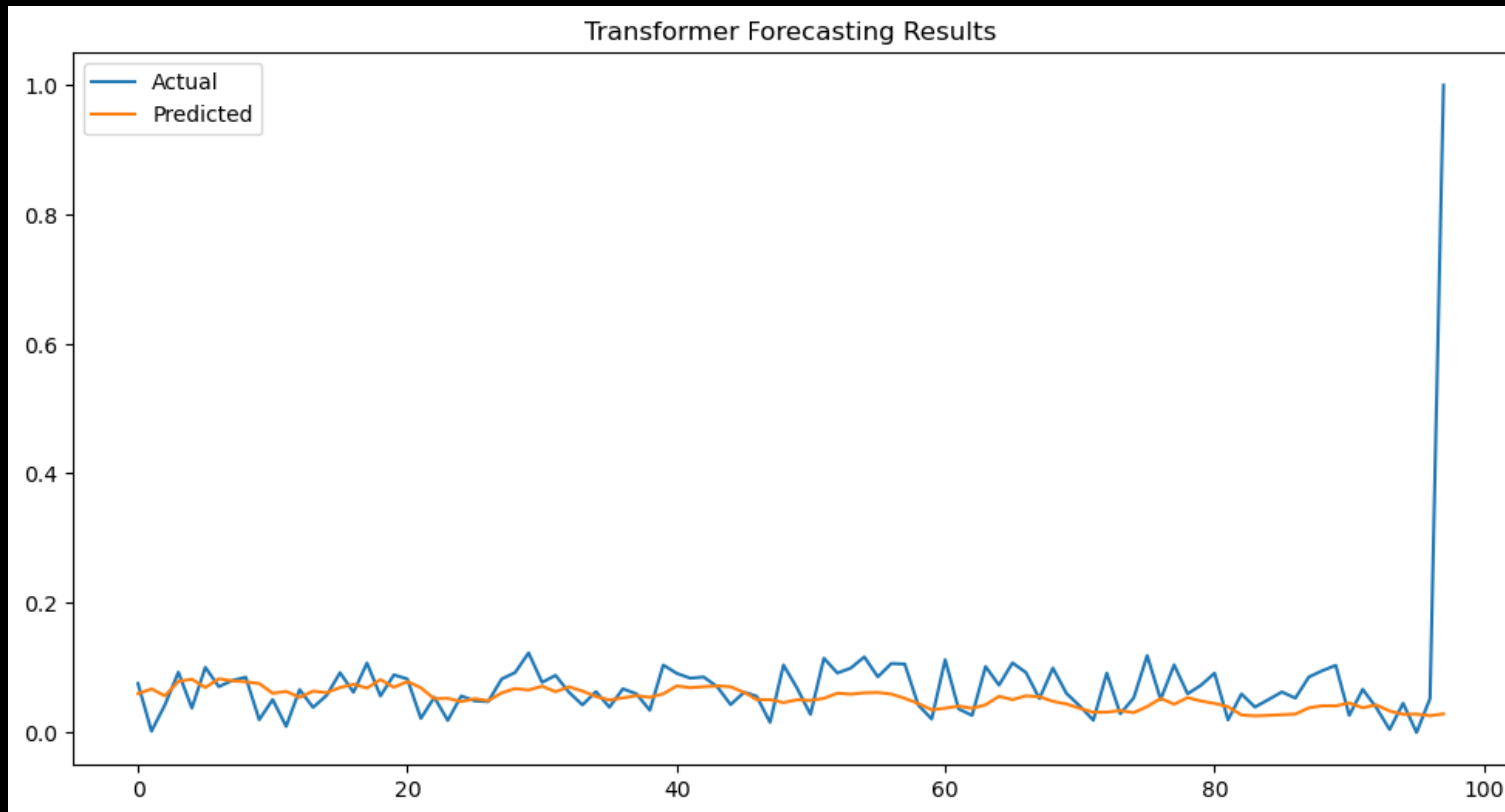
```
# Model parameters  
input_dim = X.shape[2] # Number of features  
embed_dim = 64 # Embedding dimension  
num_heads = 4 # Number of attention heads  
ff_dim = 128 # Feedforward network dimension  
num_layers = 3 # Number of transformer layers  
dropout = 0.1 # Dropout rate
```

# Optimizer and Loss Function

- `criterion = torch.nn.MSELoss()`
- `optimizer = optim.Adam(tmodel.parameters(), lr=0.001)`

```
Epoch 5/50, Loss: 0.0200
Epoch 6/50, Loss: 0.0175
Epoch 7/50, Loss: 0.0158
Epoch 8/50, Loss: 0.0138
Epoch 9/50, Loss: 0.0127
Epoch 10/50, Loss: 0.0116
Epoch 11/50, Loss: 0.0104
Epoch 12/50, Loss: 0.0095
Epoch 13/50, Loss: 0.0104
Epoch 14/50, Loss: 0.0083
Epoch 15/50, Loss: 0.0078
Epoch 16/50, Loss: 0.0084
Epoch 17/50, Loss: 0.0077
Epoch 18/50, Loss: 0.0074
Epoch 19/50, Loss: 0.0076
Epoch 20/50, Loss: 0.0070
Epoch 21/50, Loss: 0.0068
Epoch 22/50, Loss: 0.0065
Epoch 23/50, Loss: 0.0063
Epoch 24/50, Loss: 0.0060
Epoch 25/50, Loss: 0.0058
Epoch 26/50, Loss: 0.0057
Epoch 27/50, Loss: 0.0062
Epoch 28/50, Loss: 0.0058
Epoch 29/50, Loss: 0.0060
Epoch 30/50, Loss: 0.0059
Epoch 31/50, Loss: 0.0056
Epoch 32/50, Loss: 0.0054
Epoch 33/50, Loss: 0.0050
Epoch 34/50, Loss: 0.0053
Epoch 35/50, Loss: 0.0050
Epoch 36/50, Loss: 0.0050
Epoch 37/50, Loss: 0.0044
Epoch 38/50, Loss: 0.0047
Epoch 39/50, Loss: 0.0046
Epoch 40/50, Loss: 0.0049
Epoch 41/50, Loss: 0.0049
Epoch 42/50, Loss: 0.0047
Epoch 43/50, Loss: 0.0044
Epoch 44/50, Loss: 0.0044
Epoch 45/50, Loss: 0.0043
Epoch 46/50, Loss: 0.0041
Epoch 47/50, Loss: 0.0039
Epoch 48/50, Loss: 0.0046
Epoch 49/50, Loss: 0.0045
Epoch 50/50, Loss: 0.0043
```

# Training



# Predictions

## Unscaled Future Predictions:

	0	1	2	3	4	5	6	7	\
0	1741.0094	32.8152	1309.6686	31.1271	2822.1050	1405.9308	380.1586	1956.3875	
1	1886.8129	26.1994	1275.5325	24.3964	3365.1445	1713.3722	391.3523	2178.2158	
2	1960.1750	21.9284	1373.6909	21.6222	3632.3477	1988.3779	406.1049	2433.3838	
3	2003.7330	18.0578	1505.2623	18.4523	3837.4453	2326.9846	429.4685	2763.1968	
4	2068.4082	14.1441	1656.1318	16.3812	3932.5452	2686.8577	449.2589	3133.7710	
5	2093.5562	10.8822	1793.6283	14.9633	3918.9602	2940.5046	460.7080	3417.1995	
6	2113.3562	8.9785	1896.0061	13.3408	3949.0161	3109.6438	468.7467	3625.2012	
7	2117.9890	8.1316	1961.2762	12.0100	4004.7803	3216.7402	474.0439	3765.1921	
8	2123.5330	7.9407	1994.4252	11.6604	4055.2300	3270.8459	475.9704	3838.4641	
9	2125.2400	7.8247	2006.6818	11.5188	4073.0515	3291.5183	475.9724	3863.3645	
10	2125.9209	7.7744	2012.1145	11.4549	4081.2314	3300.5908	475.9526	3874.3167	
11	2126.2141	7.7519	2014.5176	11.4265	4084.8560	3304.6008	475.9412	3879.1487	

	8
0	5425.0786
1	5530.1543
2	5942.7886
3	6459.6069
4	7017.6904
5	7406.4473
6	7693.3213
7	7864.8477
8	7981.8794
9	8027.1855
10	8047.4150
11	8056.3550

## First 12 rows of the original dataset:

	Beef 3/	Veal 3/	Pork 3/	Lamb and mutton 3/	Total red meat 3/ 4/	\
0	2335.6000	2.3000	2492.6000	9.7000	4840.2000	
1	2246.1000	3.8000	2462.4000	9.6000	4721.9000	
2	2335.6000	2.3000	2492.6000	9.7000	4840.2000	
3	2168.6000	2.9000	2319.9000	10.8000	4502.1000	
4	2185.0000	3.0000	2325.7000	9.2000	4522.9000	
5	2425.8000	3.1000	2532.2000	10.3000	4971.4000	
6	2170.7000	2.9000	2221.8000	9.4000	4404.8000	
7	2256.1000	3.0000	2278.5000	9.2000	4546.8000	
8	2255.9000	3.2000	2242.5000	10.0000	4511.7000	
9	2103.6000	3.0000	2109.2000	9.1000	4225.0000	
10	2292.3000	3.4000	2269.5000	10.2000	4575.4000	
11	2269.3000	3.3000	2308.2000	10.1000	4590.9000	

	Broilers 5/	Turkey 5/	Total poultry 4/ 5/ 6/	\
0	4137.7000	415.2000	4612.7000	
1	4051.4000	435.2000	4547.6000	
2	4137.7000	415.2000	4612.7000	
3	3882.4000	376.3000	4311.2000	
4	3651.4000	394.1000	4093.1000	
5	4370.9000	493.9000	4929.6000	
6	3892.6000	413.4000	4366.2000	
7	4038.7000	435.0000	4533.2000	
8	4072.1000	433.5000	4566.2000	
9	3724.9000	408.9000	4190.7000	
10	4011.0000	446.8000	4518.9000	
11	3917.7000	450.7000	4427.5000	

	Total red meat and poultry 4/
0	9452.9000
1	9269.5000
2	9452.9000
3	8813.3000
4	8616.0000
5	9901.0000
6	8771.0000
7	9080.0000
8	9077.9000
9	8415.7000
10	9094.3000
11	9018.4000

# Discussion

- Strengths: Transformer based predictions are much easier to get working than VAE predictions.
- Limitations: It still sucks :(
- Hyperparameter Tuning
  - latent dimensionality, learning rate, sequence length, network architecture , and loss function parameters

# Conclusions

- Successfully applied time-series forecasting techniques using a transformer-based model.
- A transformer could likely be effective for strategic decision making.
- Further testing and research are needed to more fully understand my model's lack of improvement.
- Food production statistics are effective for time-series analysis.



# References

- Some understanding and techniques have been gained from Hunter Heidenreich's tutorial on VAE's which is available on their personal blog though none of their original code has been used for this project.
- Heidenreich, H. (2024, March 3). Modern Pytorch techniques for Vaes: A comprehensive tutorial. Hunter Heidenreich, Machine Learning Engineer.  
<https://hunterheidenreich.com/posts/modern-variational-autoencoder-in-pytorch/>