



# E-Commerce Customer Analysis Dashboard

This application performs comprehensive analysis on E-commerce customer data using Linear Regression to predict customer spending patterns.

Workflow: Analyze → Clean → Visualize → Train → Test → Deploy

## Step 1: ANALYZE - Dataset Exploration

✓ Dataset loaded successfully!

Total Rows	Total Columns	Missing Values	Data Types
500	8	0	2

### Dataset Preview

	Email	Address	Avatar	Avg. Session Length	Time on App	Time on Website	Length of Membership	Yearly Amount Spent
0	mstephenson@fernandez.com	835 Frank Tunnel Wrightmouth, MI 82180-9605	Violet	34.4973	12.6557	39.5777	4.0826	587.9511
1	hduke@hotmail.com	4547 Archer Common Diazchester, CA 06566-8576	DarkGreen	31.9263	11.1095	37.269	2.664	392.2049
2	pallen@yahoo.com	24645 Valerie Unions Suite 582 Cobbborough, DC 99414-7564	Bisque	33.0009	11.3303	37.1106	4.1045	487.5475
3	riverarebecca@gmail.com	1414 David Throughway Port Jason, OH 22070-1220	SaddleBrown	34.3056	13.7175	36.7213	3.1202	581.8523
4	mstephens@davidson-herman.com	14023 Rodriguez Passage Port Jacobville, PR 37242-1057	MediumAquaMarine	33.3307	12.7952	37.5367	4.4463	599.4061
5	alvareznancy@lucas.biz	645 Martha Park Apt. 611 Jeffreychester, MN 67218-7250	FloralWhite	33.871	12.0269	34.4769	5.4935	637.1024
6	katherine20@yahoo.com	68388 Reyes Lights Suite 692 Josephbury, WV 92213-0247	DarkSlateBlue	32.0216	11.3663	36.6838	4.685	521.5722
7	awatkins@yahoo.com	Unit 6538 Box 8980 DPO AP 09026-4941	Aqua	32.7391	12.352	37.3734	4.4343	549.9041
8	vchurch@walter-martinez.com	860 Lee Key West Debra, SD 97450-0495	Salmon	33.9878	13.3862	37.5345	3.2734	570.2004
9	bonnie69@lin.biz	PSC 2734, Box 5255 APO AA 98456-7482	Brown	31.9365	11.8141	37.1452	3.2028	427.1994

## Column Information

	Column	Data Type	Non-Null Count	Null Count
0	Email	object	500	0
1	Address	object	500	0
2	Avatar	object	500	0
3	Avg. Session Length	float64	500	0
4	Time on App	float64	500	0
5	Time on Website	float64	500	0
6	Length of Membership	float64	500	0
7	Yearly Amount Spent	float64	500	0

## Statistical Summary

	Avg. Session Length	Time on App	Time on Website	Length of Membership	Yearly Amount Spent
count	500	500	500	500	500
mean	33.0532	12.0525	37.0604	3.5335	499.314
std	0.9926	0.9942	1.0105	0.9993	79.3148
min	29.5324	8.5082	33.9138	0.2699	256.6706
25%	32.3418	11.3882	36.3493	2.9304	445.0383
50%	33.082	11.9832	37.0694	3.534	498.8879
75%	33.712	12.7538	37.7164	4.1265	549.3138
max	36.1397	15.127	40.0052	6.9227	765.5185

## 🧹 Step 2: CLEAN - Data Preprocessing

### Missing Values Analysis

✓ No missing values found

### Duplicate Records Analysis

✓ No duplicate records found

### Cleaned Dataset Summary

Rows After Cleaning

500

Columns

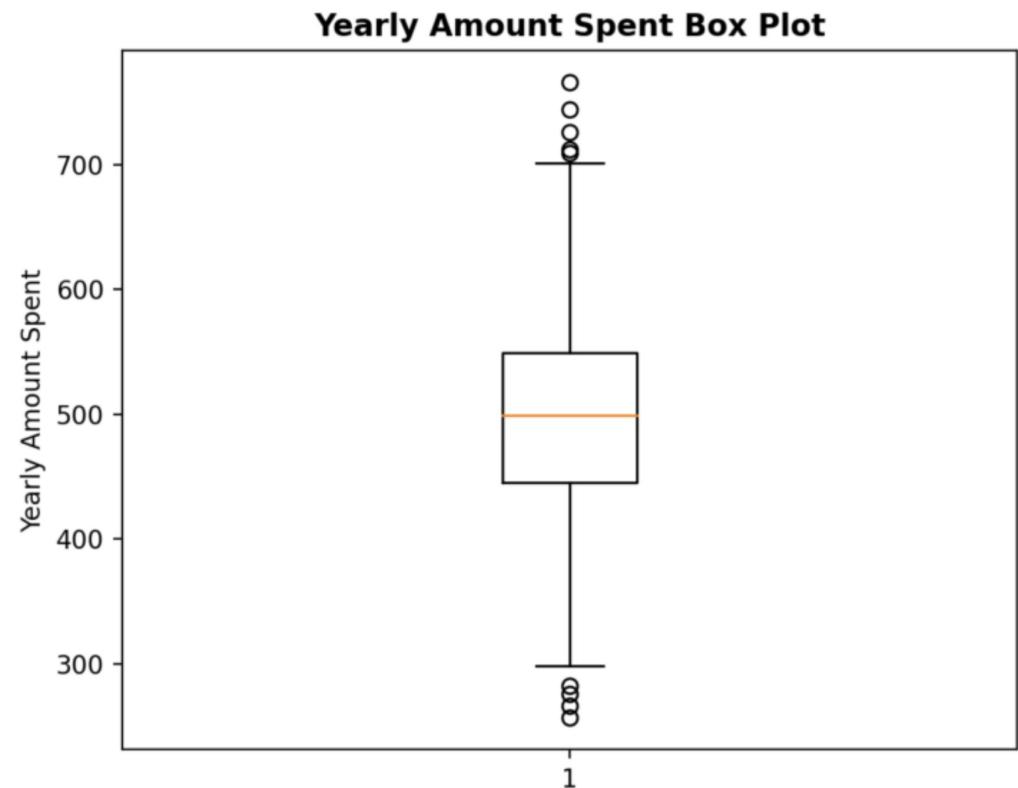
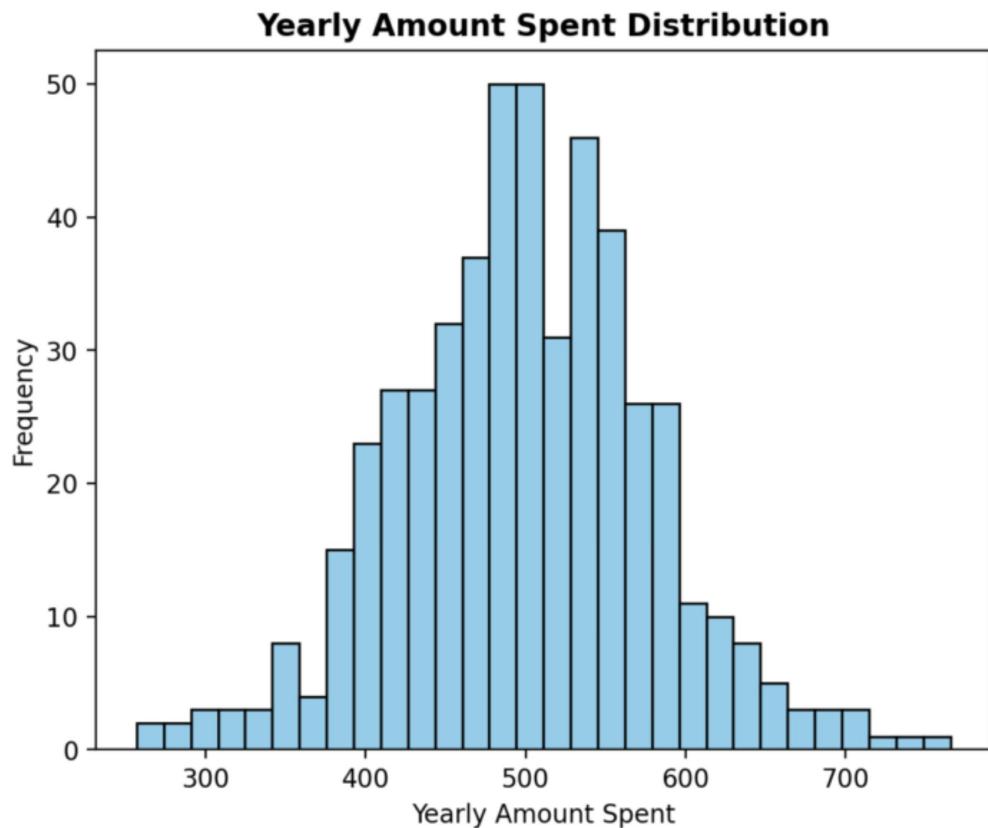
8

Rows Removed

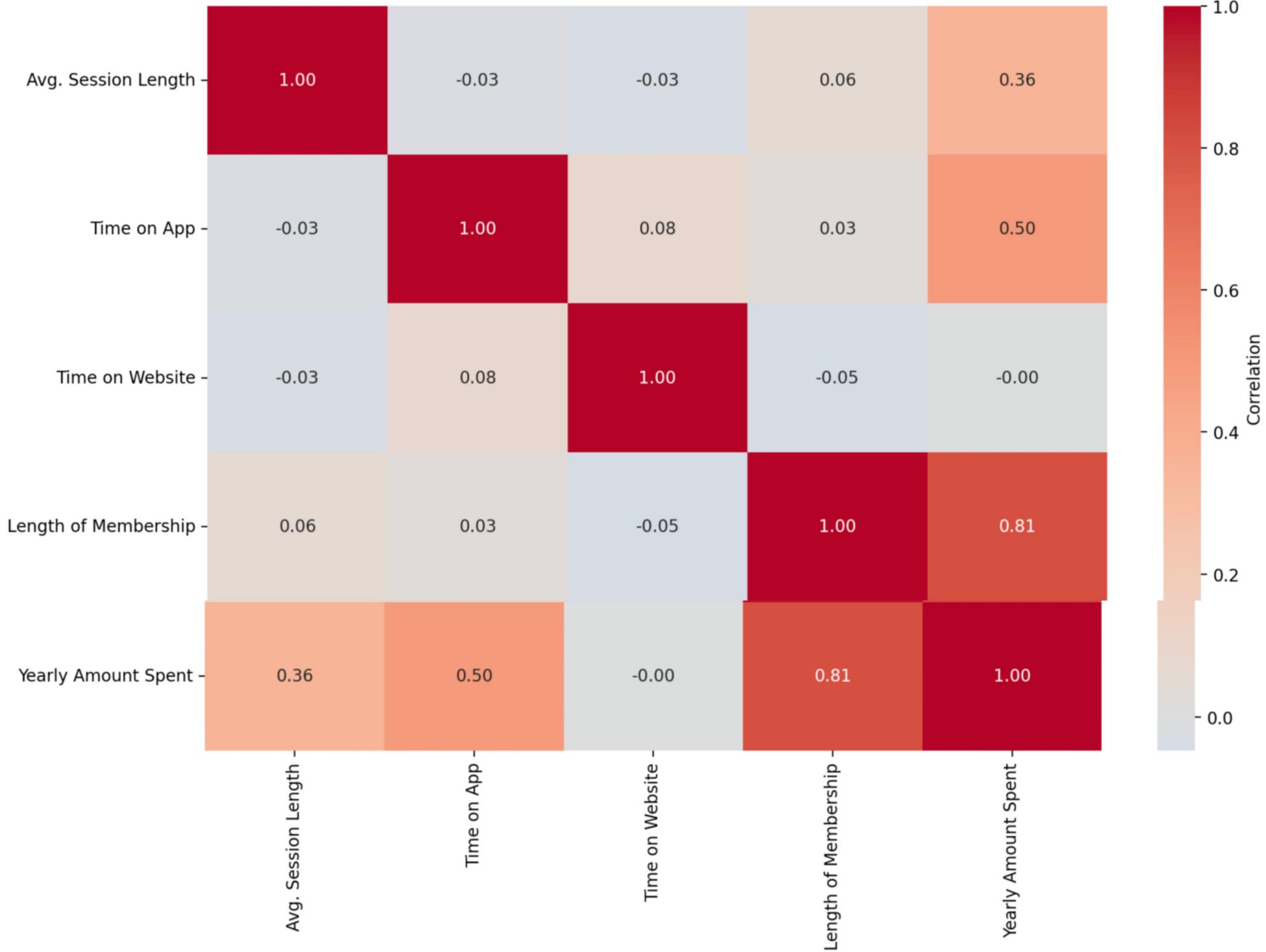
0

## Step 3: VISUALIZE - Data Exploration

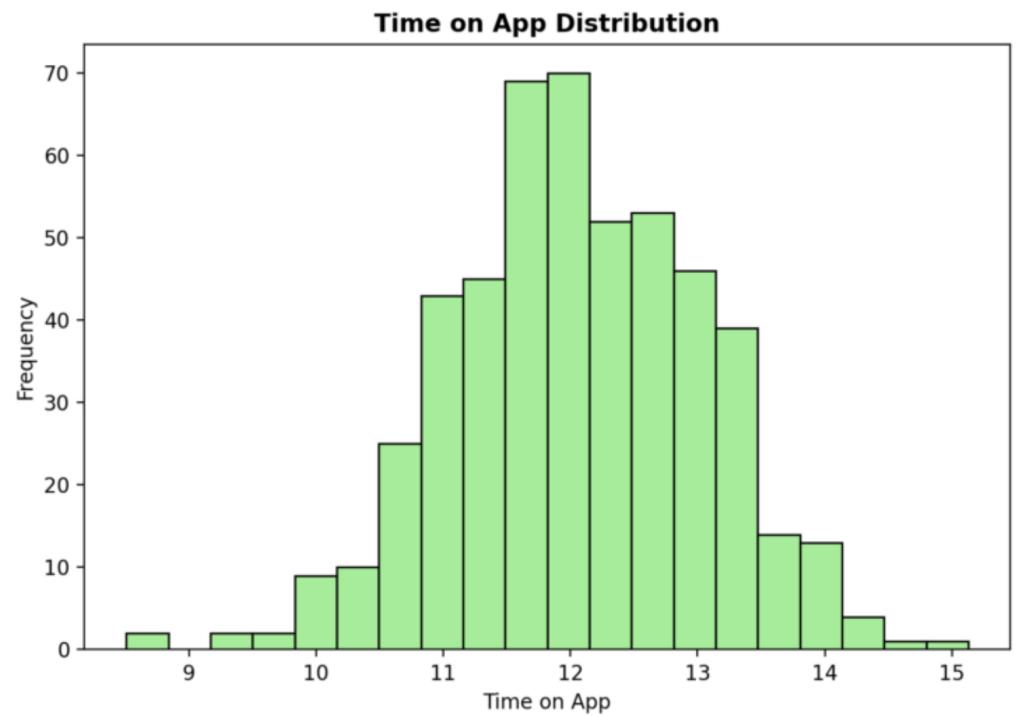
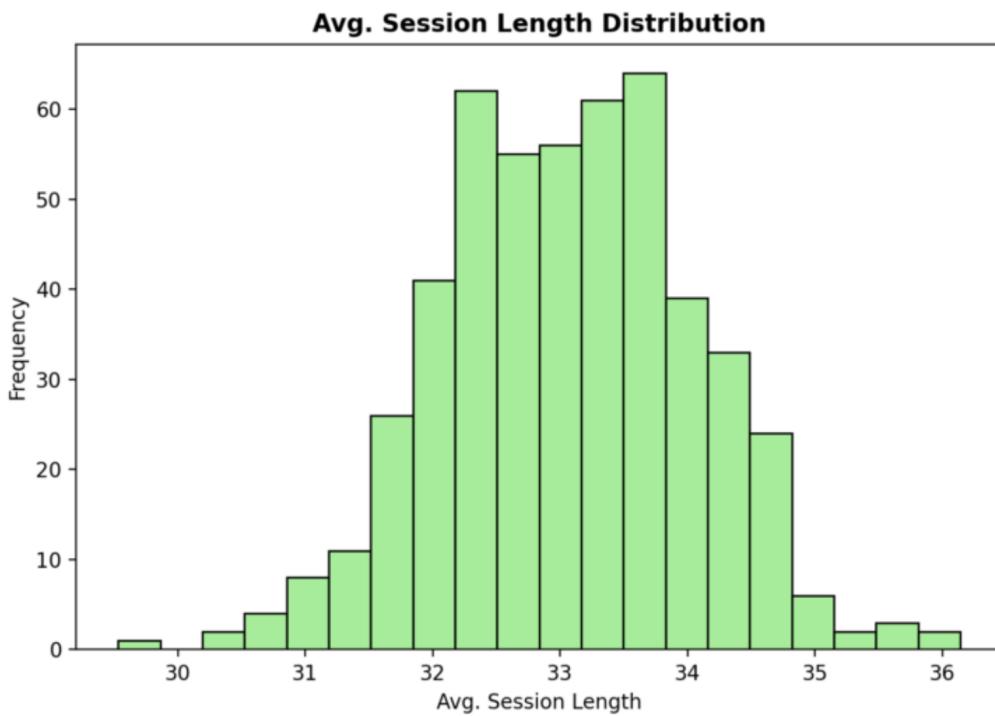
### 1. Distribution of Yearly Amount Spent

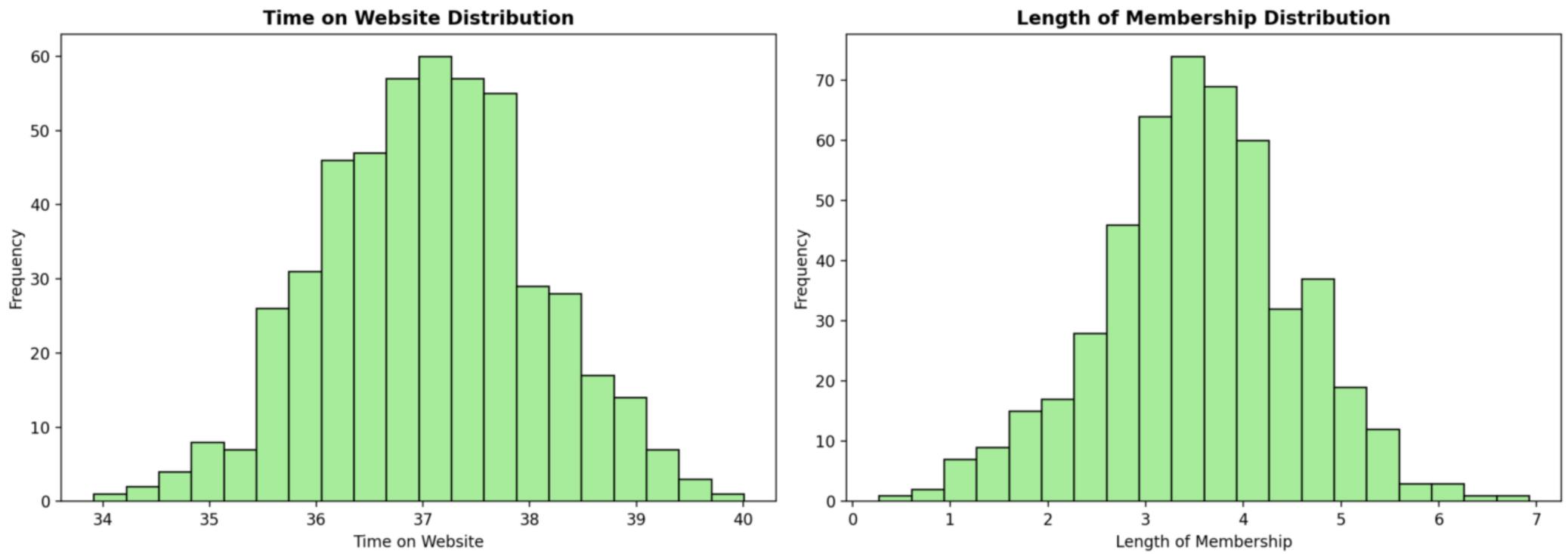


### Feature Correlation Matrix

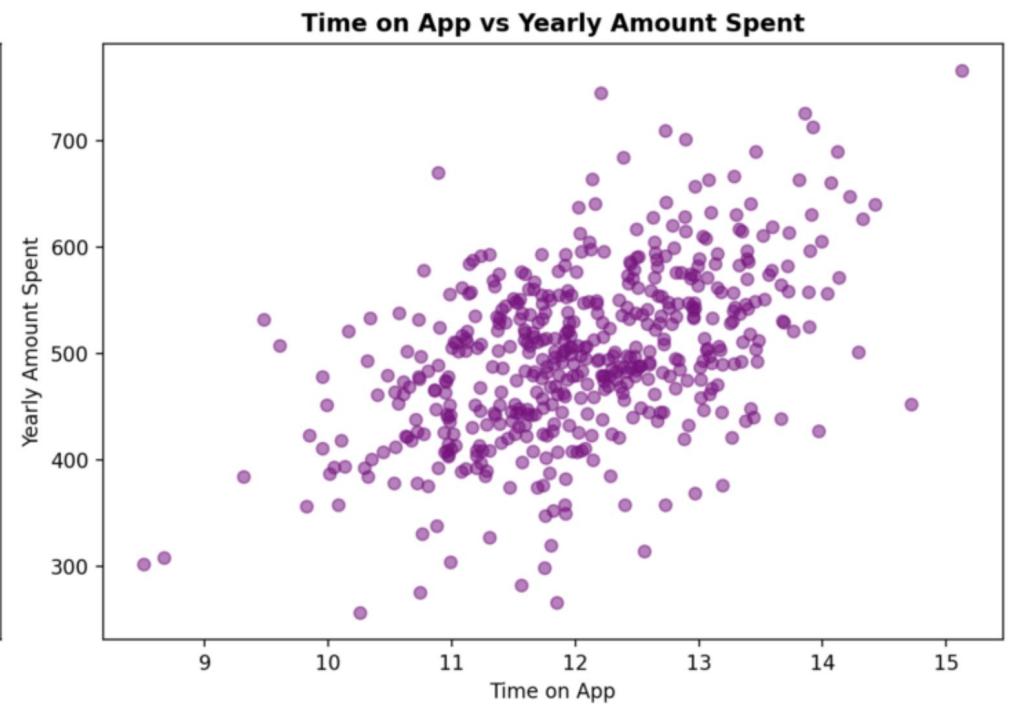
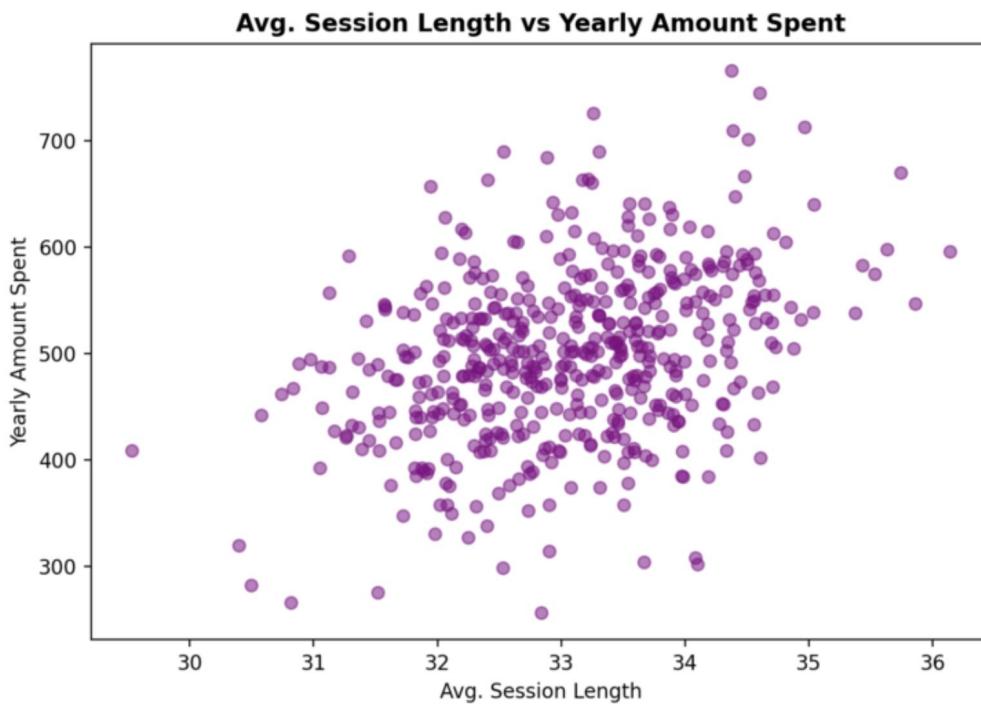


### 3. Feature Distributions

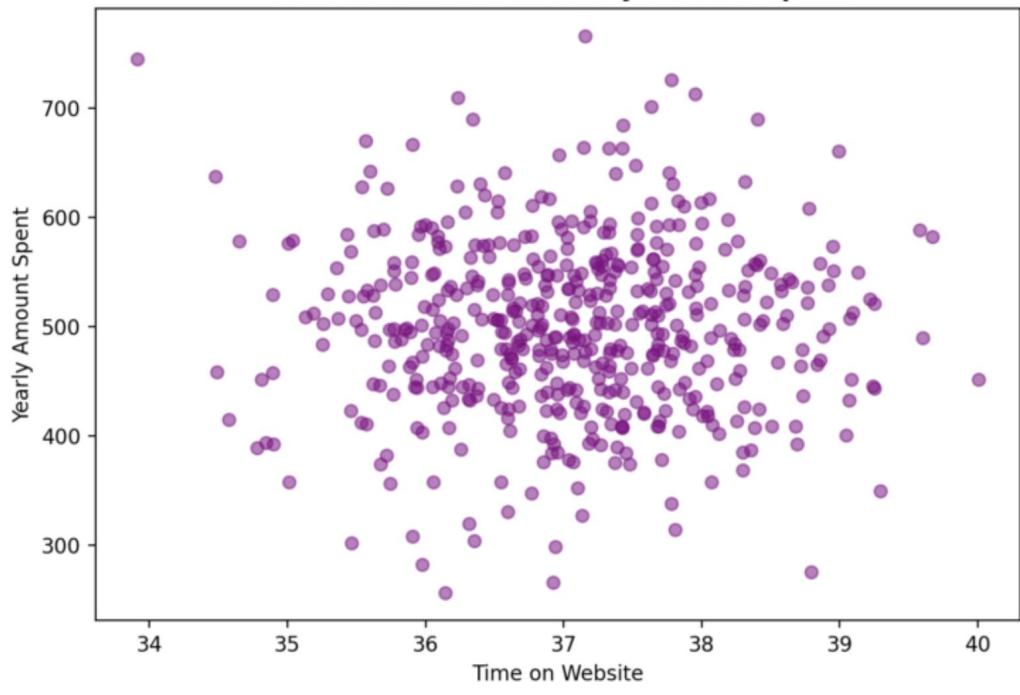




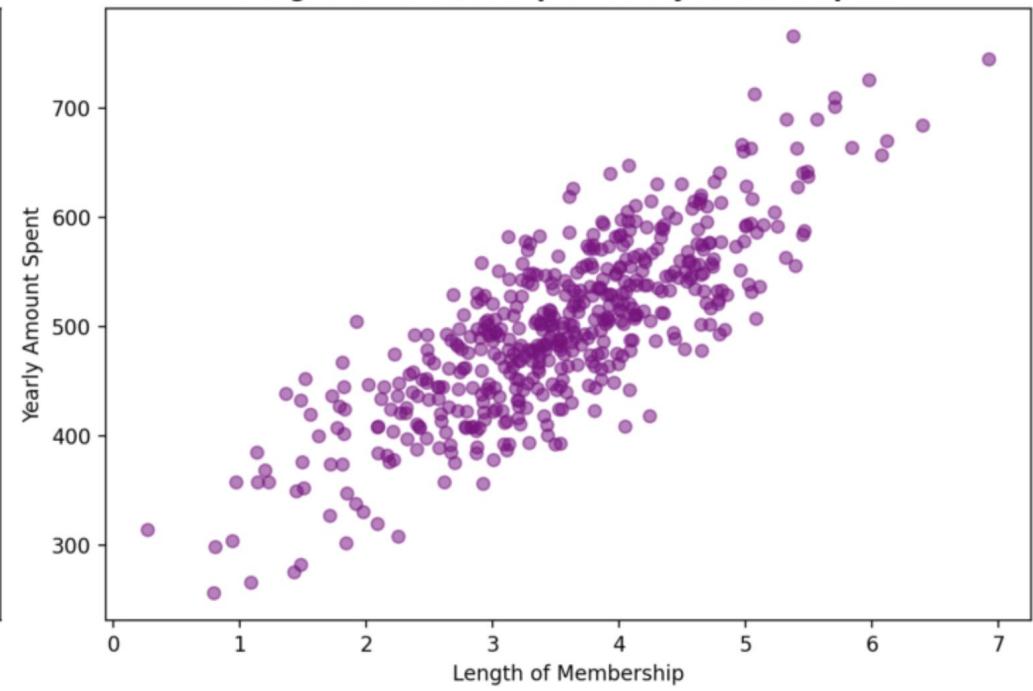
## 4. Feature vs Target Variable



**Time on Website vs Yearly Amount Spent**



**Length of Membership vs Yearly Amount Spent**



## Step 4: TRAIN - Model Development

### Data Preparation

Total Samples

500

Number of Features

4

Target Variable

Yearly Amount Spent

Selected Features:

Avg. Session Length, Time on App, Time on Website, Length of Membership

### Data Splitting (67% Train, 33% Test)

Training Samples

335

Training %

67.0%

Testing Samples

165

Testing %

33.0%

### Feature Scaling

✓ Features standardized using StandardScaler

### Model Training

✓ Linear Regression model trained successfully!

### Model Coefficients

	Feature	Coefficient	Abs_Coefficient
3	Length of Membership	60.143	60.143
1	Time on App	39.6963	39.6963
0	Avg. Session Length	26.0605	26.0605
2	Time on Website	-0.1493	0.1493

Intercept: 504.5460

## Step 5: TEST - Model Evaluation

### Performance Metrics

#### Training Metrics

MSE  
**98.6585**

RMSE  
**9.9327**

MAE  
**7.8133**

R<sup>2</sup> Score  
**0.9836 (98.36%)**

#### Testing Metrics

MSE  
**101.1398**

RMSE  
**10.0568**

MAE  
**8.1771**

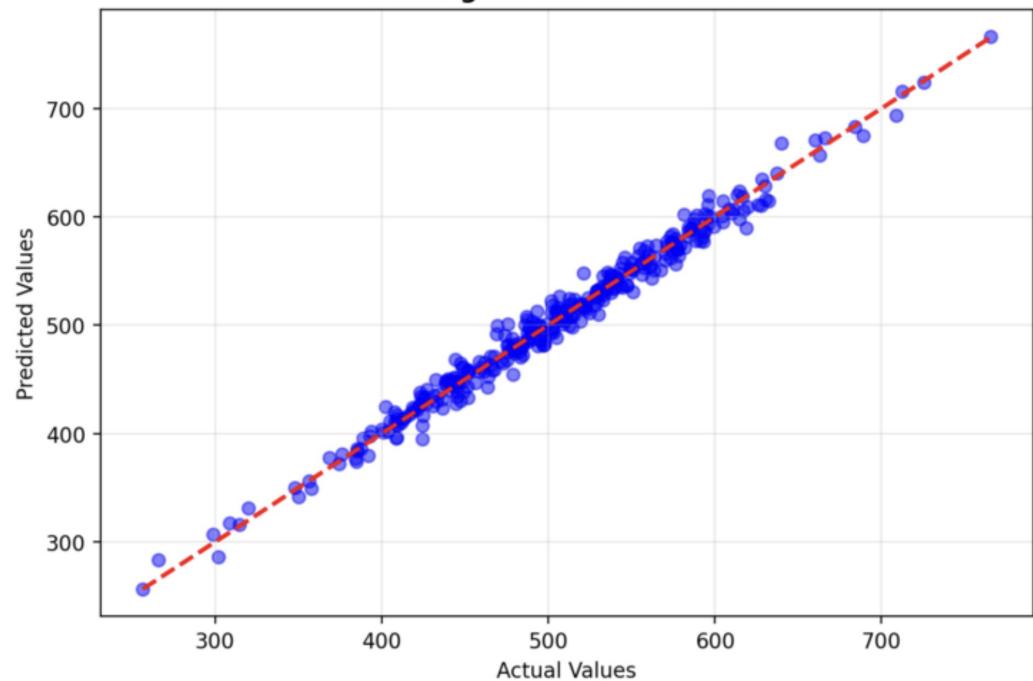
R<sup>2</sup> Score  
**0.9847 (98.47%)**

### Metrics Summary Table

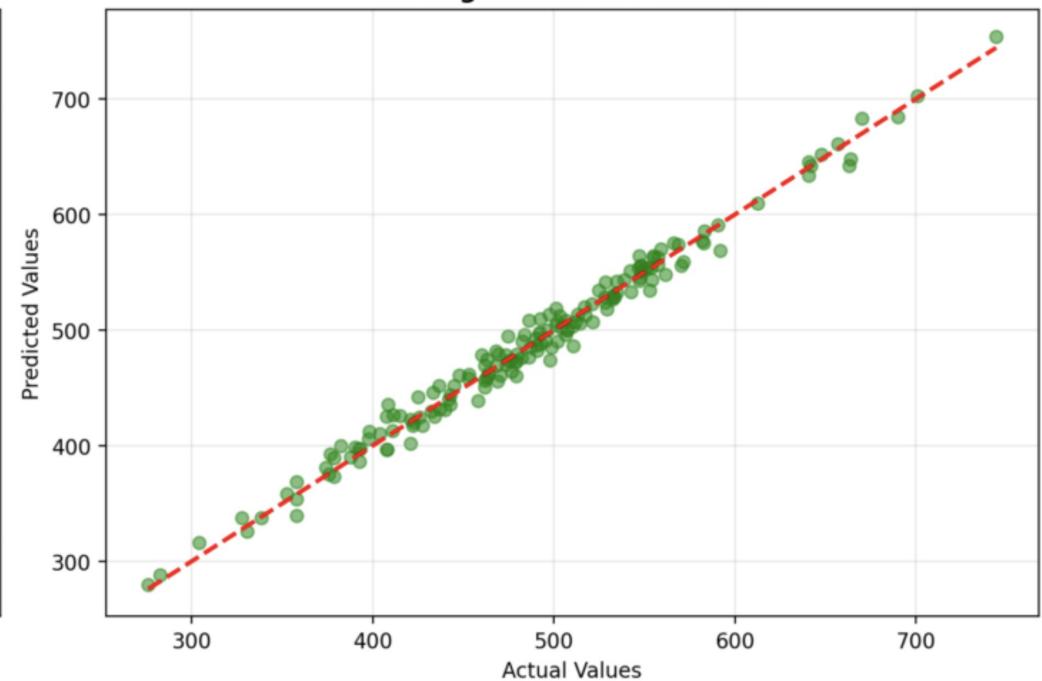
	Metric	Training	Testing
0	MSE	98.658479	101.139825
1	RMSE	9.932697	10.056830
2	MAE	7.813328	8.177101
3	R <sup>2</sup> Score	0.983638	0.984709
4	R <sup>2</sup> Score (%)	98.36%	98.47%

## Actual vs Predicted Values

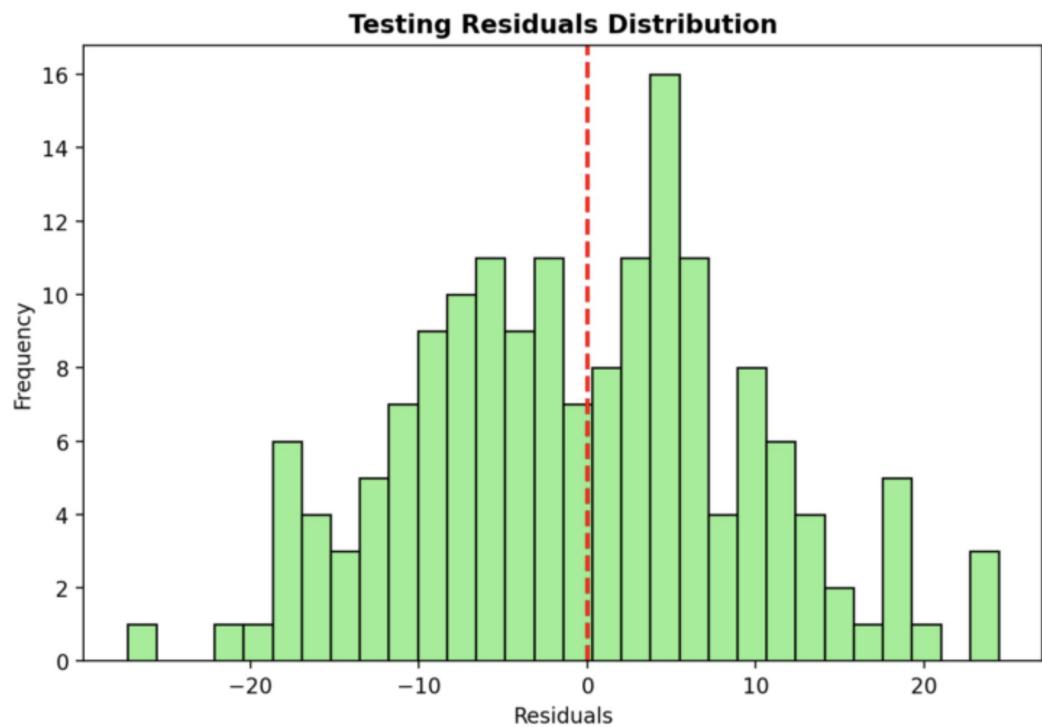
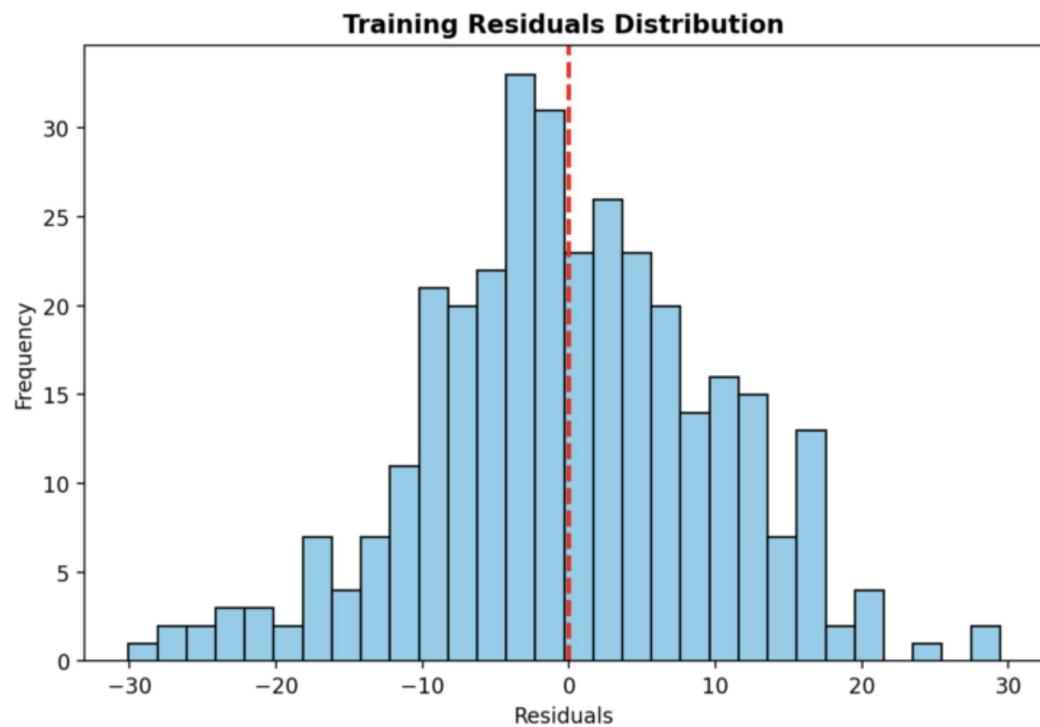
**Training: Actual vs Predicted**



**Testing: Actual vs Predicted**



## Residual Analysis



## 🚀 Step 6: DEPLOY - Make Predictions

### Predict Customer Spending

Enter customer characteristics to predict yearly spending:

Avg. Session Length	Time on App	Time on Website	Length of Membership
0.00	0.00	0.00	0.00
- +	- +	- +	- +

 **Predict Spending**

✓ Project Completed Successfully!

All workflow steps completed:

- ✓ Analyze - Dataset exploration
- ✓ Clean - Data preprocessing
- ✓ Visualize - Data visualization
- ✓ Train - Model training (67% training data)
- ✓ Test - Model evaluation
- ✓ Deploy - Prediction deployment