# ViT TEST Features Analysis Report

Generated: 2025-10-12 08:50:31

## 1. Basic Statistics

Shape: (75000, 768)

Data type: float32

Memory size: 219.7 MB

Mean: -0.0153

Std: 0.4282

Min: -0.9929

Max: 0.9922

## 2. Sample Features (First 3 Products)

Product 1:

First 10 features: [-0.63842857 -0.58757806 0.7318741 -0.32992923 -0.4011749 -0.3566763

0.6751809 -0.48786023 0.20751882 0.31919062]

Feature mean: -0.0156

Feature std: 0.4279

Product 3:

First 10 features: [-4.7768813e-01 -5.8022952e-01 5.2716655e-01 1.4786422e-01

-3.8647914e-01 -6.0273618e-01 2.4916267e-01 3.8616251e-02

-1.5306473e-04 5.0834191e-01]

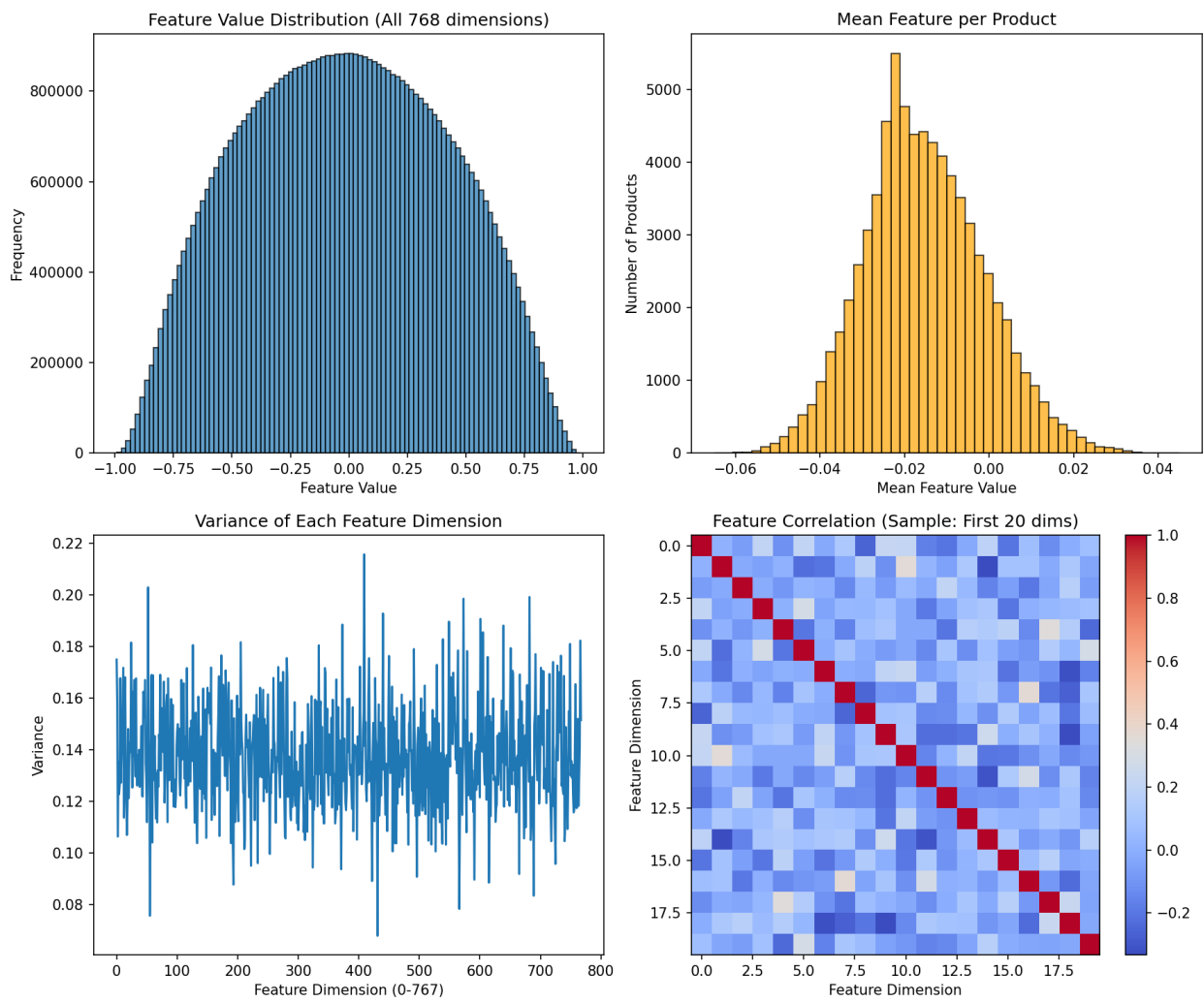Feature mean: -0.0227

Feature std: 0.4273

Product 9:

First 10 features: [-0.44975927 0.13778496 0.22360317 -0.65955967 0.36600828 -0.66497016

-0.14037445 -0.6355785 -0.19690949 0.19435006]

Feature mean: 0.0238

Feature std: 0.4209
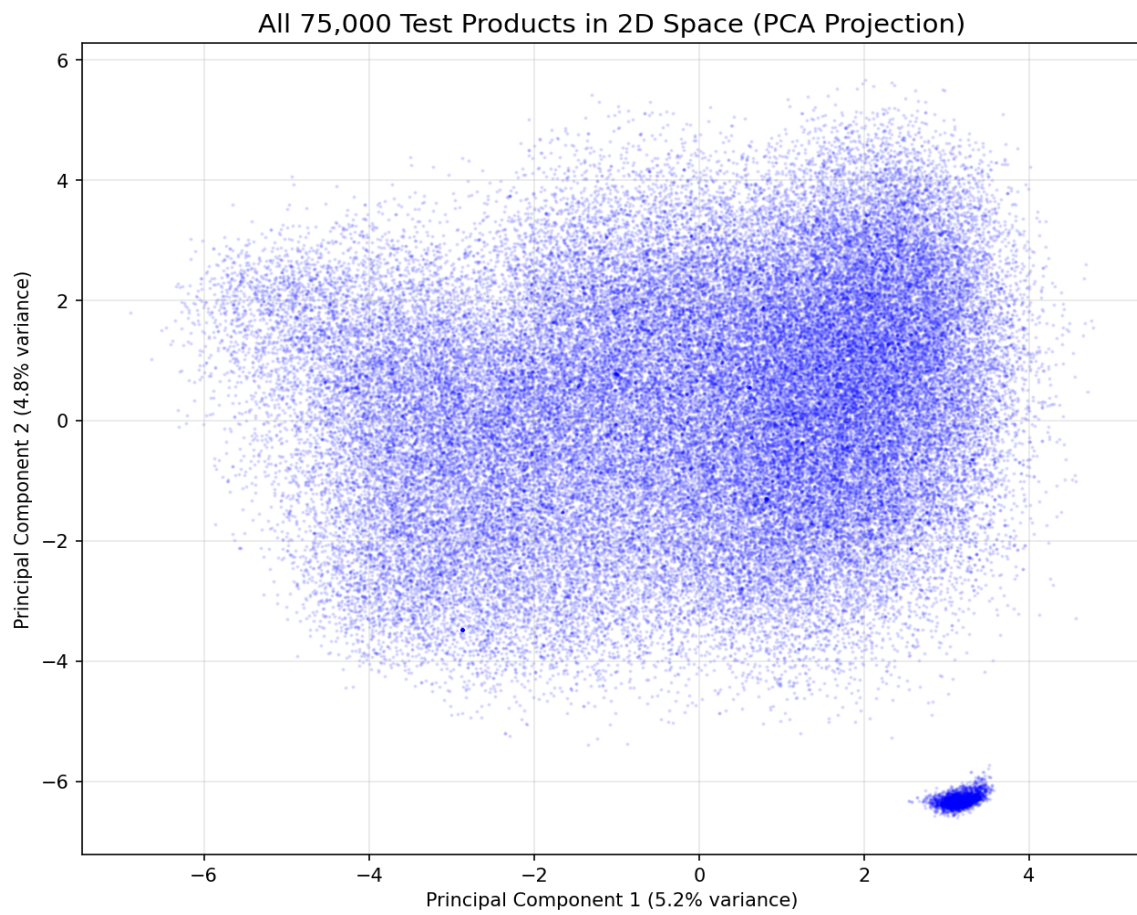
# 3. Feature Distribution Visualizations



Feature Value Distribution (All 768 dimensions)

Mean Feature per Product

Variance of Each Feature Dimension

Feature Correlation (Sample: First 20 dims)

# 4. PCA 2D Visualization

PC1 explains 5.22% of variance

PC2 explains 4.84% of variance

Total explained: 10.06%



All 75,000 Test Products in 2D Space (PCA Projection)

## 5. Metadata Analysis

```
Total products: 75,000

Successful: 74,999

Missing: 1
```

## 6. Product Similarity Analysis

Cosine similarity between first 5 products:

```
Product 1 ↔ Product 3: 0.4785 similarity

Product 1 ↔ Product 9: 0.2065 similarity

Product 1 ↔ Product 19: 0.3839 similarity

Product 1 ↔ Product 20: 0.3306 similarity

Product 3 ↔ Product 9: 0.1100 similarity

Product 3 ↔ Product 19: 0.2528 similarity

Product 3 ↔ Product 20: 0.2366 similarity

Product 9 ↔ Product 19: 0.2568 similarity

Product 9 ↔ Product 20: 0.0952 similarity

Product 19 ↔ Product 20: 0.1783 similarity
```