```
import streamlit as st
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans

# ---------------------------------------------
# APP CONFIGURATION
# ---------------------------------------------
st.set_page_config(page_title="Student Performance Analyzer", layout="wide")

# ---------------------------------------------
# APP TITLE
# ---------------------------------------------
st.title(" Student Performance Analyzer (Medical Lab Equipment)")
st.write("""
This app evaluates biomedical device readings from student lab experiments
and reports the **accuracy** and **error rates** of devices used.
Upload your CSV file to begin analysis.
""")

# ---------------------------------------------
# FILE UPLOAD
# ---------------------------------------------
uploaded_file = st.file_uploader(" Upload your CSV file", type=["csv"])

if uploaded_file:
    # Read the uploaded CSV
    df = pd.read_csv(uploaded_file)
    st.subheader(" Uploaded Data")
    st.dataframe(df)

    # Check if required columns exist
    required_cols = {"Expected Value", "Observed Value"}
    if not required_cols.issubset(df.columns):
        st.error(" CSV must include columns: 'Expected Value' and 'Observed Value'")
    else:
        # ---------------------------------------------
        # CALCULATE ERROR AND ACCURACY
        # ---------------------------------------------
        df["Error %"] = abs(df["Expected Value"] - df["Observed Value"]) / df["Expected Value"] * 100
        df["Accuracy %"] = 100 - df["Error %"]
```

```
    st.subheader(" Calculated Error and Accuracy")
    st.dataframe(df)

    # ------------------------------------------------
    # SUMMARY STATISTICS
    # ------------------------------------------------
    st.subheader(" Summary Statistics")
    st.write(df[["Expected Value", "Observed Value", "Error %", "Accuracy %"]].describe())

    # ------------------------------------------------
    # ERROR DISTRIBUTION PLOT
    # ------------------------------------------------
    st.subheader(" Error Distribution")
    fig, ax = plt.subplots()
    sns.histplot(df["Error %"], bins=10, kde=True, color="skyblue", ax=ax)
    ax.set_xlabel("Error %")
    ax.set_ylabel("Frequency")
    st.pyplot(fig)

    # ------------------------------------------------
    # KMEANS CLUSTERING
    # ------------------------------------------------
    st.subheader(" Performance Clustering (KMeans)")
    kmeans = KMeans(n_clusters=3, random_state=0)
    df["Cluster"] = kmeans.fit_predict(df[["Error %"]])

    st.bar_chart(df["Cluster"].value_counts())
    st.write(df.groupby("Cluster")[["Error %", "Accuracy %"]].mean())

    # ------------------------------------------------
    # DOWNLOAD ANALYZED RESULTS
    # ------------------------------------------------
    csv = df.to_csv(index=False).encode("utf-8")
    st.download_button(
        " Download Analyzed Data",
        data=csv,
        file_name="analyzed_results.csv",
        mime="text/csv"
    )

else:
```

```
    st.info(" Please upload a CSV file with columns 'Expected Value' and 'Observed Value' to start analysis.")
```