```
import streamlit as st import pandas as pd import numpy as np import plotly.express as px from wordcloud import WordCloud, STOPWORDS import matplotlib.pyplot as plt st.title("Sentiment Analysis Dashboard") st.sidebar.title("Sentiment Analysis") st.sidebar.markdown("Analyze the sentiment of your text data.") st.sidebar.markdown("This app is a streamlit dashboard for sentiment analysis.") DATA_URL = "Tweets.csv" @st.cache_data def load_data(): data = pd.read_csv(DATA_URL) data['tweet_created'] = pd.to_datetime(data['tweet_created']) return data data = load_data() st.sidebar.subheader("Show random tweets") random_tweet = st.sidebar.radio('Sentiment', ('positive', 'neutral', 'negative')) st.sidebar.markdown(data.query('airline_sentiment == @random_tweet') [['text']].sample(n=1).iat[0, 0]) st.sidebar.markdown("### Number of tweets by sentiment") select = st.sidebar.selectbox('Visualization type', ['Histogram', 'Pie chart'], key='1') sentiment_count = data['airline_sentiment'].value_counts() sentiment_count = pd.DataFrame({'Sentiment': sentiment_count.index, 'Tweets': sentiment_count.values}) if not st.sidebar.checkbox("Hide" , True): st.markdown("### Number of tweets by sentiment") if select == 'Histogram': fig = px.bar(sentiment_count, x='Sentiment', y='Tweets', color='Tweets', height=500) st.plotly_chart(fig) else: fig = px.pie(sentiment_count, names='Sentiment', values='Tweets') st.plotly_chart(fig) st.sidebar.subheader("When and where are users tweeting from?") hour = st.sidebar.slider("Hour of day", 0, 23) modified_data = data[data['tweet_created'].dt.hour == hour] if not st.sidebar.checkbox("Close", True , key='2'): st.markdown("### Tweets location based on the hour of day") st.markdown("%i tweets between %i:00 and %i:00" % (len(modified_data), hour, (hour + 1) % 24)) st.map(modified_data) if st.sidebar.checkbox("Show raw data", False): st.write(modified_data) st.sidebar.subheader("Breakdown by airline") choice = st.sidebar.multiselect('Pick airlines', ('US Airways', 'United', 'American', 'Southwest', 'Delta', 'Virgin America'), key='0') if len(choice) > 0: choice_data = data[data.airline.isin(choice)] fig_choice = px.histogram(choice_data, x='airline', y='airline_sentiment', histfunc='count', color='airline_sentiment', facet_col='airline_sentiment', labels={'airline_sentiment':'tweets'}, height=600, width=800) st.plotly_chart(fig_choice) st.sidebar.subheader("Word Cloud") word_sentiment = st.sidebar.radio('Display word cloud for what sentiment?', ('positive', 'neutral', 'negative')) if not st.sidebar.checkbox("Close", True , key='3'): st.header('Word cloud for %s sentiment' % (word_sentiment)) df = data[data.airline_sentiment == word_sentiment] words = ' '.join(df['text']) processed_words = ' '.join([word for word in words.split() if 'http' not in word and not word.startswith('@') and word != 'RT']) wordcloud = WordCloud(stopwords=STOPWORDS , background_color='white').generate(processed_words) plt.imshow(wordcloud, interpolation='bilinear') plt.xticks([]) plt.yticks([]) st.pyplot(plt)
```