Case Study: Weather Data Analysis Let us look at a case study of using the big data stack for 1.9 analysis of weather data.To collect and ingest streaming sensor data generated by the weather monitoring stations, we can use a publish-subscribe messaging framework to ingest data for real-time analysis within the Big Data stack and a Source-Sink connector to ingest data into a distributed filesystem for batch analysis.Let us say, we want our weather analysis application to aggregate data on various timescales (minute, hourly, daily or monthly) to determine the mean, maximum and minimum readings for temperature, humidity, wind and pressure.To collect and ingest streaming sensor data generated by the weather monitoring stations, we can use a publish-subscribe messaging frameworkData Preparation Since the weather data received from different monitoring stations can have missing values, use different units and have different formats, we may need to prepare data for analysis by cleaning, wrangling, normalizing and filtering the data.Data Collection Let us assume, we have multiple weather monitoring stations or end-nodes equipped with temperature, humidity, wind, and pressure sensors.We also want the application to support interactive querying for exploring the data, for example, queries such as: finding the day with the lowest temperature in each month of a year, finding the top-10 most wet days in the year, for instance.Next, we also want the application to make predictions of certain weather events, for example, predict the occurrence of fog or haze.Analysis Modes Based on the analysis types determined the previous step, we know that the analysis modes required for the application will be batch, real-time and interactive.Visualizations The front end application for visualizing the analysis results would be dynamic and interactive.Mapping Analysis Flow to Big Data Stack Now that we have the analytics flow for the application, let us map the selections at each step of the flow to the big data stack.To come up with a selection of the tools and frameworks from the big data stack that can be used for weather data analysis, let us first come up with the analytics flow for the .application as shown in Figure 1.10.These type of analysis come under the basic statistics category

lakhasly.com © تم تلخيص النص بواسطة موقع لخصلي