Solutions:

1. An advantage is the added potential complexity that we can represent using that new layer; if we are using image data, the convolutional layer can provide valuable spatial information. At the same time, we risk violating Occam's Razor, which states that we ought to choose the simplest model with the same efficacy.
2. This is known as the Dense layer and the parameters for it include the number of units and an optional activation function.
3. The Conv2D layer takes the number of filters, kernel size, an optional activation function, a padding scheme, and a stride. The only required parameters are the number of filters and kernel size. All other parameters are optional and default to a reasonable value.
4. Spark provides a nice interface through the map-reduce paradigm of programming. It also provides a variety of Pandas-esque data operations.
5. Text datasets such as Shakespeare's texts would be very well suited to Spark's parallelism.
6. True. This is explained in the note in the advanced features of Keras where it discusses the data generators.
7. True. Most convolutional neural networks tend to use RELU throughout the network except for the final (couple of) activation(s).
8. False. It provides many machine learning models as discussed in the notes as well.
9. True. Pandas only operates in singlethreaded mode and thus cannot exploit multicore systems whereas Spark can.
10. The function can be implemented as follows:

```
def evaluate_mean_cent(model_loc, X_val, y_val):
      model = tf.keras.models.load_model(model_loc)
      loss_val = 0
      y_pred = model.predict(X_val - X_val.mean())
      loss_val = loss(y_pred, y_val - y_val.mean())
      return loss_val
```