



— Lion's Den —  
**RISK MODELLING  
CHALLENGE**  
Warsaw | 31 May - 1 June 2019

Please, solve the below pre-selection assignment and send your result to [lionsden@ing.com](mailto:lionsden@ing.com) until April 26.

In a particular Polish bank branch, the amount of daily cash withdrawals was analysed. This type of analysis helps optimize the amount of cash stored at the branch, as excess cash signifies a measurable loss for the bank. The first stage of this analysis consisted of collecting data on cash withdrawals made on business days for the five following quarters: Q1 2018 to Q1 2019. We would like you to undertake the second stage of analysis, the details of which you can find below.

We have specified a set of analyses that will help you get acquainted with the nuances of the issue at hand and to detect patterns within the data. Your analysis can be carried out in many ways and the results don't have to be clear cut. Remember, that your method of obtaining the result is equally important to the result itself. We have listed out a set of analysis below, but you're not required to perform all of them.

### Your Task

The following data is available:

- *date* – daily date [DD.MM.YYYY];
- *working* – variable that takes the YES value when the given day is a working day, NO otherwise;
- *withdrawal* – the sum of cash withdrawals on a given day [in thousand PLN];

[The data set in the \*withdrawals.csv\* file is here>>](#)

1. Calculate:
  - a. descriptive statistics: arithmetic mean, standard deviation, quartiles, skewness coefficient, kurtosis,
  - b. the withdrawals histogram,
  - c. outliers. What could be the cause of increased demand for cash and subsequent withdrawals?
2. Verify if the data on withdrawals follows normal distribution using statistical tests.

Use your results from points 1-2 and complete 3-6:

3. Assess the impact of the **day of the week** on the size of withdrawals. Is there a relationship between these variables and how can these dependencies be justified?
4. Assess the impact of the **day of the month** on the size of withdrawals. Is there a relationship between these variables and how can such dependencies be justified?
5. Assuming that withdrawals follow normal distribution, what is the probability that withdrawals from a given day are in the range from PLN 220,000 to PLN 250,000? Can this be confirmed by the historical data?
6. Find the long-term trend of the withdrawals and justify the hypothesis that an increasing trend exists.