# Data Science for Manufacturing: Data set radial forging at AFRC testbed v2

Radial forging is widely used in industry to manufacture components for a broad range of sectors including automotive, medical, aerospace, rail and industrial. The Advanced Forming Research Centre (AFRC) houses a GFM SKK10/R radial forge that has been used to generate a dataset archived[1] at:

https://zenodo.org/record/3405265#.Yjnk3-fP2Hs

The process use two pairs of "hammers" operating at 1200 strokes/min, and providing a maximum forging force per hammer of 150 tons, the radial forge is capable of processing a range of metals, including steel, titanium and inconel. Both hollow and solid material can be formed with the added benefit of creating internal features on hollow parts using a mandrel. Parts can be formed at a range of temperatures from ambient temperature to 1200°C.The AFRC's radial forging process was used to produce a total of 81 parts over one day of operation. The process of manufacturing each of the 81 parts was recorded by 99 sensors. The dataset contains .csv files named from "Scope0001" to "Scope0081" that each recorded the production of an individual part.
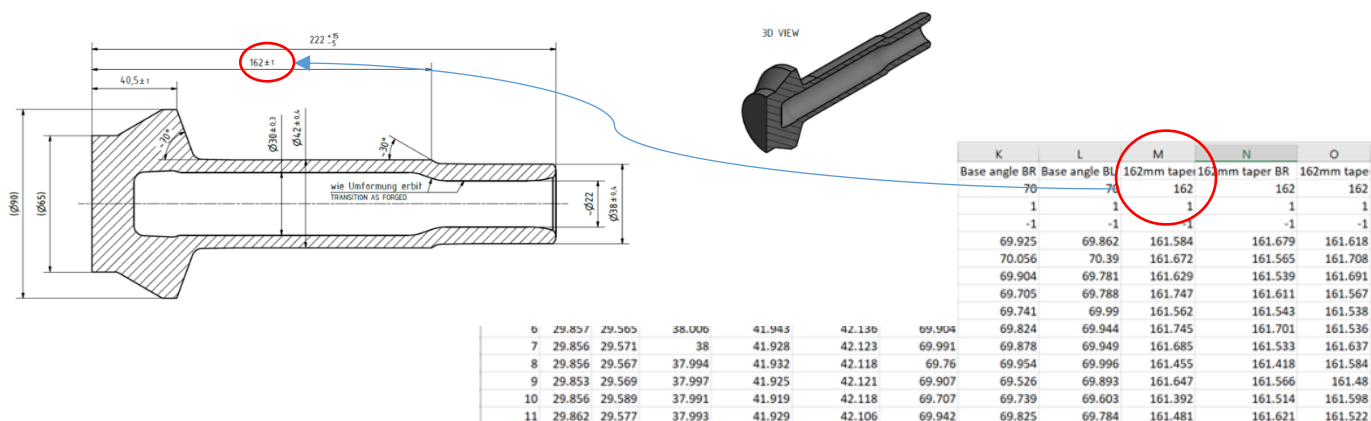
Each forged part was then measured using a Coordinate Measuring Machine (CMM) that recorded the size of 16 key dimensions (and their nominal specification and tolerances) for each of 81 parts forged. Review this dataset and:

0) Describe the data cleansing and preparation steps require to enable the following tasks.

1) Propose, and justify, a combined metric (i.e. a single none binary value) that characterises the quality of the part based solely on the CMM data.

2) Rank the sensors you judge most important in order of their influence on your quality metric.

3) For the 5 highest ranked sensors illustrate how their variance correlates with your quality metric.

4) Discuss the results and suggest which 3 process parameters (i.e. sensor values) should be most tightly controlled to ensure high quality parts.

For the provided data set, a total of 81 parts were forged over one day of operation. A machine failure occurred during the forging of part number 70, and this part was re-run once the malfunction had been fixed. Each forged part was then measured using a CMM to provide dimensional output relative to a target specification and tolerances. The CMM records 18 dimensional measurements.

The aim of the measurement setup is to predict the quality (in terms of dimensional properties) of the forged part from the sensor measurements during the forging process.

---

[1] The archive contains links to a number of python notebooks used to implement various Machine Learning procedures on the data, however note that none of these relate to the tasks required for this assignment.