**Dataset “Energia”**

Ogni file contiene 7 colonne, tra le quali quelle di interesse sono:

POD: identificativo del contatore di energia;

DATA: giorno di rilevazione;

ORA: valore progressivo che indica l'orario di rilevazione. Le rilevazioni sono ogni 15 minuti a partire dalle 00:00 del primo giorno del mese;

CONSUMO\_ATTIVA\_PRELEVATA: valore dei kw nei 15 minuti considerati (potenza). E' la variabile di interesse.

I dataset si prestano a diverse linee di indagine (si tratta solo di spunti generali; potete immaginare anche altre linee):

* Previsioni (su diversi orizzonti temporali e per quantità puntuali o aggregate, come l’energia assorbita complessivamente in una certa settimana o mese…)
* Effetto covid
* Confronto e clusterizzazione fra giorni o settimane, o mesi…
* Confronto fra i due edifici
* Identificazione di picchi di assorbimento (o di “non-assorbimento” anomali)
* Impatto del meteo o di altri eventi (convegni?) sui consumi.

Potete anche provare a progettare un’architettura di informazioni interessanti per un energy manager (attenzione, non un’architettura di dati…ma di informazioni significative) e progettare una dashboard.

**Attenzione**: andate a vedere la definizione di Kw e di Kwh e state attenti alla differenza fra potenza ed energia e a quale di queste due quantità può essere aggregata e a quale no…

**Nota**. I dati sono ìn formato diverso e alcuni mesi sono ripetuti in doversi file. Dovete fare un piccolo lavoro di pulizia e ricostruzione. Se vi sono problemi tecnici con l’apertura dei file, fatemelo sapere che chiedo a chi me li ha forniti.