L’Elaboratore del Collaudo

progetto formativo

stanciu nicolas

2025

Sommario

[-Obiettivi: 2](#_Toc200707859)

[-Come realizzarlo 2](#_Toc200707860)

[PARTE 1: Realizzata in C++ e Python 2](#_Toc200707861)

[PARTE 2: Realizzata in Python 2](#_Toc200707862)

[-Realizzazione pratica del lavoro 3](#_Toc200707863)

[PARTE 1: Realizzata in C++ 3](#_Toc200707864)

[-Librerie usate 3](#_Toc200707865)

[-Correzione nome del file 3](#_Toc200707866)

[-Separazione campi di ogni riga del file 3](#_Toc200707867)

[-Suddivisione valori ed identificativi dei dati 4](#_Toc200707868)

[-Preparazione dei dati alla conversione 4](#_Toc200707869)

[-Conversione dei dati 4](#_Toc200707870)

[-Riscrittura dei dati con le loro unità di misura 5](#_Toc200707871)

[-Spostamento file finale 5](#_Toc200707872)

[-Esempio del file CSV risultante 6](#_Toc200707873)

[PARTE 1: Rifatta in Python 6](#_Toc200707874)

[-Librerie usate 6](#_Toc200707875)

[-Apertura e lettura dal file 7](#_Toc200707876)

[-Conversione dati 7](#_Toc200707877)

[-Riscrittura dei dati con le loro unità di misura 7](#_Toc200707878)

[PARTE 2 (versione precedente): Realizzata in Python 8](#_Toc200707879)

[-Librerie usate 8](#_Toc200707880)

[-Definizione dei profili di stile delle celle 8](#_Toc200707881)

[-Definizione della struttura e dei contenuti delle righe 8](#_Toc200707882)

[-Gestione altezza riga ed unione celle 9](#_Toc200707883)

[-Assegnazione contenuti e applicazione stili alle celle 9](#_Toc200707884)

[-Estrazione dati dal nome del file 9](#_Toc200707885)

[-Estrazione dati dal file 9](#_Toc200707886)

[-Estrazione degli elementi che compongono le versioni 10](#_Toc200707887)

[-Ordinamento degli elementi che compongono le versioni 10](#_Toc200707888)

[-Uniformazione dei nomi dei campi 10](#_Toc200707889)

[-Controllo dei campi per l’inserimento dei dati 11](#_Toc200707890)

[-Inserimento dati nel template XLSX 11](#_Toc200707891)

[PARTE 2 (versione ottimizzata): Realizzata in Python 12](#_Toc200707892)

[-Salvataggio indirizzi delle celle unite 12](#_Toc200707893)

[-Inserimento dati nel template XLSX 12](#_Toc200707894)

[-Esempio del file XLSX risultante 13](#_Toc200707895)

# -Obiettivi:

L’obiettivo del progetto è processare i dati grezzi provenienti dal collaudo, inserirli in un file CSV e successivamente trasferirli in un nuovo template XLSX, opportunamente formattato e impaginato. I dati analizzati vengono sovrascritti nel file CSV di origine, insieme al relativo identificativo. Grazie a questo identificativo, il programma può collocare i dati nelle celle corrette del file XLSX. Il file finale viene poi spostato nella cartella dedicata al cliente. Il programma deve inoltre essere in grado di gestire dati provenienti da più file CSV diversi, inserendoli automaticamente nel template XLSX.

# -Come realizzarlo

Per realizzare il progetto è necessario suddividerlo in due fasi e seguire i passaggi elencati per entrambe

## PARTE 1: Realizzata in C++ e Python

1 - Eseguire uno script PS1 per ricavare più file CSV contenenti i dati rilevati da un collaudo

2 – Controllare la correttezza dei nomi dei file e se non coerenti modificarli

3 – Analizzare ogni file riga per riga estraendo i dati necessari ed inserendoli in una struttura adeguata

4 – Convertire i dati fuori scala in modo da renderli più leggibili

5 – Scorrere la struttura e riscrivere i dati contenuti, insieme alle relative unità di misura, all’interno di un nuovo file CSV

6 – Una volta che i dati sono stati convertiti ed il file CSV è stato formattato, il file viene spostato in una cartella dedicata, così da renderlo disponibile per la seconda fase del progetto

## PARTE 2: Realizzata in Python

1 – Definire una struttura per memorizzare i profili di stile dei testi

2 – Definire una struttura che contenga i dettagli di ogni riga: altezza, celle da unire, contenuto e profilo di stile

3 – Analizzare e modificare la struttura delle righe secondo i parametri indicati. Il template xlsx finale deve essere poi spostato in una cartella dedicata.

4 – Memorizzare in una struttura adeguata i dati contenuti nel file csv formattato

5 – Ricavare eventuali dati non presenti direttamente nel CSV o separati all’interno di esso

6 – Analizzare gli identificativi presenti nel template XLSX ed i campi nella struttura per trovare le corrispondenze ed inserire i dati nelle rispettive celle

7 – Salvare il file finale all’interno di una cartella dedicata

# -Realizzazione pratica del lavoro

## PARTE 1: Realizzata in C++

### -Librerie usate

Per la realizzazione del programma in C++ sono state utilizzate le seguenti librerie:

- “iostream”: per la gestione dell’input/output da console

- “filesystem”: per la gestione dei file e delle cartelle

- “fstream”: per la lettura e la scrittura dei file CSV.

- “vector”: per la gestione di array dinamici, utilizzati per memorizzare dati e tipi di dati letti dai file.

- “map”: per associare i nomi dei campi alle rispettive unità di misura.

### -Correzione nome del file

for (int index{0}; index < filename.length(); index++) {

     if (filename[index] == '.') {

         num\_dots++;

     }

     if (!std::isalnum(filename[index]) && filename[index] != '\_') {

         if (filename[index] == '.' && num\_dots <= 1) continue;

filename[index] = '\_';

}

Tramite un ciclo FOR, ogni carattere del nome del file viene analizzato. Se il ciclo rileva caratteri diversi da lettere e numeri, questi vengono sostituiti con un carattere underscore (“\_”). Inoltre, se nel nome del file è presente più di un punto, anche i punti successivi al primo vengono sostituiti con underscore.

### -Separazione campi di ogni riga del file

for (int index{0}; index <= line.length(); index++) {

     if (index == line.length() || line[index] == separator) {

         line\_parts.push\_back(line.substr(start\_index, index - start\_index));

         start\_index = index + 1;

     }

}

push\_back = metodo utilizzato per aggiungere un elemento alla fine di un vettore. Accetta come parametro un elemento dello stesso tipo del vettore su cui viene chiamato.

substr = metodo che estrae una parte della stringa, dato l’indice di inizio e la lunghezza del segmento da estrarre.

Per estrarre le parole da ogni campo di una riga, un ciclo FOR analizza i caratteri e, al rilevamento del separatore o della fine riga, estrae il segmento tra i separatori o tra l’ultimo separatore e la fine.

### -Suddivisione valori ed identificativi dei dati

if (line\_parts.at(0) == "TIME") {

    for (std::string str : line\_parts) {

         data\_type.push\_back(str);

    }

} else {

    for (std::string str : line\_parts) {

         data.push\_back(str);

    }

}

Viene verificata la prima parola dell’array che contiene i campi della riga. Se questa parola è uguale a “TIME”, significa che la riga rappresenta gli identificativi dei campi e non i dati grezzi. In questo modo, è possibile memorizzare separatamente gli identificativi in un array e i dati grezzi in un altro.

### -Preparazione dei dati alla conversione

for (int data\_field{0}; data\_field < data.size(); data\_field++) {

     std::string new\_data{ data.at(data\_field) };

     while (new\_data.find('.') != -1) {

           new\_data.erase(new\_data.find('.'), 1);

     }

     data.at(data\_field) = new\_data;

}

find = metodo che cerca una sottostringa o un carattere all’interno della stringa e restituisce l’indice della prima occorrenza, oppure -1 se non viene trovato.

erase = metodo che rimuove una parte della stringa, specificando l’indice di inizio e la lunghezza del segmento da eliminare.

Tutti i dati memorizzati vengono analizzati tramite un ciclo FOR e, per ciascun dato, vengono rimossi i punti. Questo passaggio facilita le successive operazioni di conversione sui dati.

### -Conversione dei dati

for (int data\_field{0}; data\_field < data.size(); data\_field++) {

     for (auto field : fields\_to\_modify){

          if (data\_type.at(data\_field).find(field) != -1){

              double value = std::stod(data.at(data\_field)) \* 0.000001;

              std::string new\_data = std::to\_string(value);

              if (new\_data.find('.') != -1){

                  new\_data = new\_data.substr(0, new\_data.find('.') + 2);

              }

              data.at(data\_field) = new\_data;

           }

      }

}

I dati raccolti vengono esaminati scorrendo la struttura in cui sono memorizzati. I valori che necessitano di conversione vengono trasformati in double, successivamente moltiplicati per 0,000001 e infine riconvertiti in stringa.

### -Riscrittura dei dati con le loro unità di misura

for (int data\_field{0}; data\_field < data\_type.size(); data\_field++) {

     line.clear();

     line += data\_type.at(data\_field) + ";";

     line += data.at(data\_field) + ";";

     for (auto mu : measurement\_units){

        if (data\_type.at(data\_field).find(mu.first) != -1) line += mu.second + ";"

     }

     if (data\_field == data\_type.size() - 1) {

         output\_file << line;

     } else {

         output\_file << line << std::endl;

     }

}

clear = metodo che elimina tutti i caratteri dalla stringa, rendendola vuota.

Dopo aver ottenuto e convertito tutti i dati, il programma procede a riscriverli nel file CSV originale. Scorrendo una mappa che associa a ciascun tipo di dato la relativa unità di misura, il programma collega ogni dato estratto dalla struttura ai rispettivi identificativi e, se previsto, assegna loro l’unità di misura. In questo modo vengono create stringhe che includono l’identificativo del campo, il valore e l’unità di misura, che vengono poi scritte all’interno del file CSV.

### -Spostamento file finale

bool moveFile(std::string filename, std::string destination){

    try {

        std::filesystem::create\_directory(destination);

        std::filesystem::rename(filename, destination + "/" + filename);

    } catch (const std::filesystem::filesystem\_error& e) {

        return false;

    }

    return true;

}

create\_directory =  metodo che crea una cartella nel percorso specificato dal path fornito.

rename = metodo che consente di rinominare o spostare un file specificando il nuovo nome o percorso.

Si avvia una funzione che riceve come parametri il nome del file e la cartella in cui lo si vuole inserire. Viene utilizzato un try-catch per poter gestire l’eccezione lanciata nel caso in cui non si riesca ad effettuare lo spostamento.

### 

### -Esempio del file CSV risultante



## PARTE 1: Rifatta in Python

La parte 1, precedentemente realizzata in C++, è stata ora riscritta in Python. Le funzionalità implementate e il risultato finale rimangono invariati rispetto alla versione precedente.

### -Librerie usate

Per la realizzazione del programma in Python sono state usate le seguenti librerie:

- “pathlib”: utilizzata per la gestione dei percorsi di file e cartelle

- “csv”: utilizzata per la lettura e la scrittura dei file CSV

- “os”: utilizzata per operazioni sul filesystem, come la creazione di cartelle

### -Apertura e lettura dal file

with open(current\_file.name, newline = '', encoding = 'utf-8') as csvfile:

      reader = csv.reader(csvfile, delimiter = ',')

      for row in reader:

          if len(row) > 0:

              if row[0].lower() == 'time':

                fields\_name = row

              elif len(fields\_data) == 0:

                fields\_data = row

open = metodo usato per aprire un file e restituire un oggetto file. Accetta come parametri il percorso, la modalità, l’encoding e altre opzioni.

reader =  funzione che legge un file CSV e restituisce ogni riga come una lista di valori, separando i campi in base al delimitatore specificato

Dopo aver aperto il file CSV, il programma suddivide ogni riga utilizzando il delimitatore. Se la riga contiene almeno un elemento, verifica se si tratta dell’intestazione dei campi o dei dati. Una volta distinti i due tipi di riga, i relativi valori vengono memorizzati in una struttura dati.

### -Conversione dati

for key, value in csv\_data.items():

     for field\_to\_modify in fields\_to\_modify:

          if field\_to\_modify in key.lower():

              csv\_data[key] = str(float(value))

              fields\_modified = True

              break

Alcuni dati vengono convertiti in float per essere riportati su una scala più leggibile e successivamente riconvertiti in stringa. Questa operazione viene eseguita identificando i campi da convertire tramite un array che contiene i nomi dei parametri interessati.

### -Riscrittura dei dati con le loro unità di misura

for field\_name, field\_data in csv\_data.items():

     measurement\_unit = ''

     for type\_unit, unit in measurement\_units.items():

          if type\_unit in field\_name.lower():

              measurement\_unit = unit

              field\_name = field\_name.replace('\_Mevo', '')

              writer.writerow([field\_name, field\_data, measurement\_unit])

writerow = metodo che scrive una riga nel file CSV, accettando una lista di valori da inserire come campi nella riga.

Scorrendo un dizionario che associa a ciascun tipo di dato la relativa unità di misura, il programma collega ogni dato estratto dalla struttura ai rispettivi campi e assegna l’unità di misura appropriata, se prevista. Ogni riga del file finale sarà quindi composta dal nome del campo, dal valore del dato e dalla sua unità di misura.

## PARTE 2 (versione precedente): Realizzata in Python

### -Librerie usate

Per la realizzazione del programma in Python sono state usate le seguenti librerie:

* “openpyxl”:
* “PIL”:
* “pathlib”:
* -“re”:

### -Definizione dei profili di stile delle celle

cells\_profiles = {

   0: {

      'alignment':Alignment(wrap\_text=True, horizontal='left, vertical='center'),

      'font': Font(size = 7.5),

       'border': Border(left = thin, top = thin, right = thin, bottom = thin),

       'fill': PatternFill()

      },

Viene creato un dizionario in cui ogni chiave rappresenta l’ID di un profilo, mentre il valore associato è un altro dizionario. Quest’ultimo contiene, come chiavi, i nomi dei parametri da modificare e, come valori, le relative impostazioni o modifiche da applicare.

### -Definizione della struttura e dei contenuti delle righe

cells\_layout = []

cells\_layout.append({'ROW\_HEIGHT': [15], 'B-N': ['DATI GENERALI', 5]})

cells\_layout.append({'ROW\_HEIGHT': [25], 'B-G': ['', 12], 'H-N': ['', 12]})

Viene creata una lista in cui ogni elemento è un dizionario. In ciascun dizionario, la chiave rappresenta un identificativo che indica il tipo di operazione da eseguire: ad esempio, ‘ROW\_HEIGHT’ specifica che deve essere modificata l’altezza della riga, mentre una chiave nel formato ‘X-Y’ indica le colonne da unire. Il valore associato alla chiave contiene i dettagli dell’operazione: per ‘ROW\_HEIGHT’ è indicata la nuova altezza della riga, mentre per ‘X-Y’ sono specificati il contenuto della cella e il profilo di stile da applicare.

### -Gestione altezza riga ed unione celle

if key == 'ROW\_HEIGHT':

 ws.row\_dimensions[cell\_row].height = value[0]

else:

 start\_merge\_cell\_col = key.split('-')[0] if '-' in key else key

 end\_merge\_cell\_col = key.split('-')[1] if '-' in key else key

 ws.merge\_cells(f'{start\_merge\_cell\_col}{cell\_row}:{end\_merge\_cell\_col}{cell\_row}')

ws = oggetto che rappresenta il foglio XLSX di lavoro

merge\_cells = metodo che unisce più celle, accettando come parametro l’intervallo (riga e colonna iniziale e finale) da unire.

Ogni chiave del dizionario viene analizzata e, in base al suo valore, si procede a modificare l’altezza della riga oppure a unire più celle. Per effettuare l’unione, dalla chiave nel formato ‘X-Y’ vengono estratte le colonne iniziale e finale, che identificano l’intervallo delle celle da unire.

### -Assegnazione contenuti e applicazione stili alle celle

ws[cell\_address] = value[0]

ws[cell\_address].alignment = cells\_profiles[value[1]]['alignment']

ws[cell\_address].font = cells\_profiles[value[1]]['font']

ws[cell\_address].fill = cells\_profiles[value[1]]['fill']

Il contenuto viene assegnato alle celle e, in base al profilo associato, vengono applicati gli stili di formattazione corrispondenti.

### -Estrazione dati dal nome del file

csv\_data = {}

info\_parts = current\_filename.split('\_')

csv\_data['Matricola'] = info\_parts[1]

csv\_data['Commessa'] = info\_parts[0]

csv\_data['Conferma'] = info\_parts[1][:5]

Le informazioni contenute nel nome del file CSV formattato vengono estratte e aggiunte, con il rispettivo identificativo, al dizionario che raccoglie tutti i dati.

### -Estrazione dati dal file

for line in current\_csv:

     line\_parts = line.split(';')

     if line\_parts[0].lower() == 'time':

         line\_parts[0] = 'Data'

         line\_parts[1] = line\_parts[1][:line\_parts[1].find('T')]

csv\_data[line\_parts[0].replace('\"', '')]= line\_parts[1] + ' ' + line\_parts[2]

I dati con le unità di misura contenuti nel file CSV formattato vengono estratti ed aggiunti, con il loro rispettivo identificativo, al dizionario che raccoglie tutti i dati. Dal campo ‘”time” viene estratta la data.

### -Estrazione degli elementi che compongono le versioni

versions = {}

for key, value in csv\_data.items():

    if key.find('Versione') != -1:

      version\_order = ''.join(ch if ch.isdigit() else '' for ch in key)

    versions.setdefault(key.replace(version\_order,''),{})[version\_order]=value

keys\_to\_delete = [key for key in csv\_data if 'Versione' in key]

for key in keys\_to\_delete : csv\_data.pop(key, None)

Scorrendo il dizionario che contiene tutti i dati, le chiavi che includono la parola "Versione" vengono individuate. Da ciascuna chiave viene estratto il numero che ne indica l’ordine, e la coppia (numero, valore) viene inserita come valore in un dizionario dedicato alle versioni, che ha come chiave il nome della versione senza il numero. Successivamente, tutte le chiavi contenenti "Versione" vengono rimosse dal dizionario principale dei dati

### -Ordinamento degli elementi che compongono le versioni

for key, value in versions.items():

     version\_size = max(int(size) for size in value)

     versions\_correct\_order = [None] \* version\_size

     version\_num = ''

     for version\_order, version\_value in value.items():

         versions\_correct\_order[int(version\_order) - 1] = version\_value

     for num\_version in versions\_correct\_order:

         if num\_version : version\_num += num\_version  + '.'

     version\_num = version\_num[:-1]

     csv\_data[key] = version\_num

Per ogni gruppo di versioni, i valori vengono ordinati in base al numero estratto e concatenati in un’unica stringa separata da punti, che viene poi reinserita nel dizionario principale con la chiave corrispondente al tipo di versione.

### -Uniformazione dei nomi dei campi

key\_name = key

key\_name = ''.join(ch if ch.isalnum() else ' ' for ch in key\_name)

for index in range(0, len(key\_name)):

     if key\_name[index].isupper() and index != 0 and key\_name[index - 1].islower():

         key\_name = key\_name[:index] + ' ' + key\_name[index:]

key\_name = ''.join(ch.lower() if ch.isalpha() else ch for ch in key\_name)

field\_to\_search = key\_name.split(' ')

Ogni chiave del dizionario dei dati viene normalizzata: tutti i caratteri non alfanumerici sono sostituiti da uno spazio, le parole in formato camel case vengono separate, tutte le lettere sono convertite in minuscolo e le parole risultanti vengono inserite in un array.

### -Controllo dei campi per l’inserimento dei dati

if is\_first\_in\_merged\_cell(ws, current\_cell\_address):

  for field\_part in field\_to\_search:

    field\_name = ws[current\_cell\_address].value.lower() if ws[current\_cell\_address].value else ''

    if not(ws[current\_cell\_address].value and re.search(rf'\b{field\_part}', field\_name) != None):

      can\_put\_data = False

      break

else:

    continue

Se la cella analizzata corrisponde alla prima cella di un intervallo di celle unite, si verifica che ogni parola del campo in esame sia presente nel campo corrispondente del template; solo se tutte le parole sono contenute, si autorizza l’inserimento del dato.

### -Inserimento dati nel template XLSX

if can\_put\_data is True:

for next\_col in range(col + 1, ord(template\_end\_col) - 65):

       cell\_address\_to\_put\_data = f'{chr(next\_col + 66)}{row}'

if is\_first\_in\_merged\_cell(ws, cell\_address\_to\_put\_data):

          if ws[cell\_address\_to\_put\_data].value is None:

            ws[cell\_address\_to\_put\_data] = value

              break

          elif re.search(r':\s\*$', ws[cell\_address\_to\_put\_data].value) :

                ws[cell\_address\_to\_put\_data] = ws[cell\_address\_to\_put\_data].value + ' ' + value

                break

Scorrendo le colonne successive a quella del campo trovato, se la cella è la prima di un gruppo di celle unite ed è vuota oppure contiene una parola che termina con i due punti, il dato viene inserito in quella cella.

## PARTE 2 (versione ottimizzata): Realizzata in Python

Per ridurre i tempi di esecuzione del programma, ho ottimizzato la procedura di scrittura nel template XLSX. Ora il template viene scansionato non più cella per cella, ma direttamente da una cella unita all’altra. Inoltre, sono stati aggiunti controlli per escludere automaticamente le celle non rilevanti.

### -Salvataggio indirizzi delle celle unite

for merged\_range in ws.merged\_cells.ranges:

   merged\_range\_parts = str(merged\_range).split(':')

   if len(merged\_range\_parts) < 2: merged\_range\_parts.append(merged\_range\_parts[0])

   merged\_cells\_ranges[merged\_range\_parts[0]] = merged\_range\_parts[1]

Durante la scansione delle celle unite nel foglio Excel, per ogni intervallo unito viene salvato in un dizionario l’indirizzo della prima cella come chiave e quello dell’ultima cella dell’intervallo come valore. Questo permette di identificare rapidamente i limiti di ciascuna area unita nel foglio di lavoro.

### -Inserimento dati nel template XLSX

next\_merged\_cell\_address = chr(ord(last\_merged\_col\_address[0]) + 1) + last\_merged\_col\_address[1:]

while ord(next\_merged\_cell\_address[0]) < ord(template\_end\_col):

       if not(ws[next\_merged\_cell\_address].value):

           ws[next\_merged\_cell\_address] = value

           break

       elif re.search(r':\s\*$', ws[next\_merged\_cell\_address].value):

              ws[next\_merged\_cell\_address] = ws[next\_merged\_cell\_address].value + ' ' + value

              break

       next\_merged\_cell\_col = chr(ord(merged\_cells\_ranges[next\_merged\_cell\_address][0]) + 1)

       next\_merged\_cell\_row = str(merged\_cells\_ranges[next\_merged\_cell\_address][1:])

       next\_merged\_cell\_address =  next\_merged\_cell\_col + next\_merged\_cell\_row

Dopo aver individuato la fine di una cella unita, il codice calcola l’indirizzo della cella immediatamente successiva sulla stessa riga. Entra quindi in un ciclo che scorre verso destra, controllando ogni cella fino al limite della tabella (definito da template\_end\_col).  
Se trova una cella vuota, vi inserisce il valore desiderato e interrompe il ciclo.  
Se invece la cella contiene già un valore che termina con i due punti (:), aggiunge il nuovo valore in coda, separandolo con uno spazio, e poi esce dal ciclo.  
Se nessuna delle condizioni precedenti è soddisfatta, aggiorna l’indirizzo della cella successiva da controllare, spostandosi ancora a destra, utilizzando le informazioni sulle celle unite memorizzate nella mappa. In questo modo, il codice garantisce che i dati vengano inseriti nella prima cella disponibile dopo un intervallo di celle unite, oppure concatenati correttamente se necessario.

### -Esempio del file XLSX risultante

