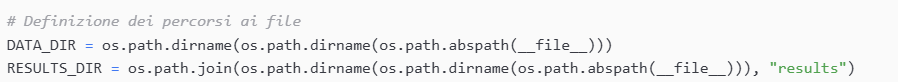
**1. Intestazione del file e importazioni**

A screen shot of a computer

AI-generated content may be incorrect.

* La prima riga (#!/usr/bin/env python3) indica che lo script deve essere eseguito con l'interprete Python 3.
* Il commento # -\*- coding: utf-8 -\*- specifica che il file è codificato in UTF-8.
* La stringa di documentazione (""" ... """) descrive brevemente lo scopo dello script.
* Le importazioni includono moduli standard (os, pandas, argparse) e funzioni personalizzate (load\_sequence, get\_protein\_sequence, plot\_variant\_distribution, plot\_variant\_types, generate\_summary\_report) dai moduli sequence\_utils e visualize.
* **2. Definizione delle directory**



* DATA\_DIR e RESULTS\_DIR definiscono i percorsi delle directory per i dati di input e i risultati, rispettivamente.os.path.dirname(os.path.dirname(os.path.abspath(\_\_file\_\_))) si riferisce alla directory principale del progetto.
* **3. Parsing degli argomenti della riga di comando**

A computer code with text

AI-generated content may be incorrect.

* La funzione parse\_arguments utilizza il modulo argparse per definire e analizzare gli argomenti della riga di comando:
  + --gene: il gene da analizzare.
  + --sequence\_file: il file FASTA contenente la sequenza genetica.
  + --variants\_file: il file CSV contenente le varianti genetiche.
* **4. Caricamento delle varianti**

A computer code with green and black text

AI-generated content may be incorrect.

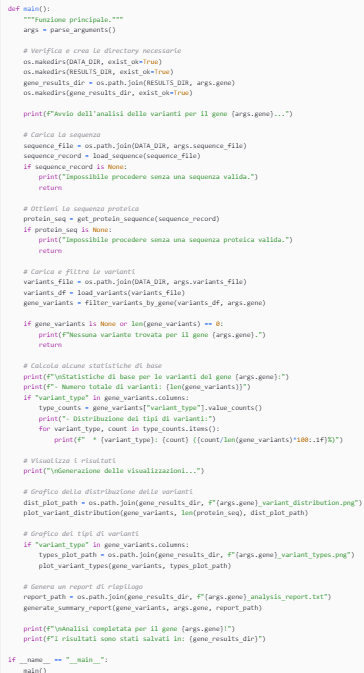
* La funzione load\_variants legge un file CSV contenente varianti genetiche e lo carica in un DataFrame di pandas. Se c'è un errore durante il caricamento, viene stampato un messaggio di errore e la funzione restituisce None.

**5. Filtraggio delle varianti per gene**

A screen shot of a computer code

AI-generated content may be incorrect.

* La funzione filter\_variants\_by\_gene filtra le varianti nel DataFrame per un particolare gene (gene\_id). Restituisce un DataFrame contenente solo le varianti del gene specificato.
* **6. Funzione principale**



* La funzione main esegue i seguenti passaggi:
  1. Analizza gli argomenti della riga di comando.
  2. Verifica e crea le directory necessarie.
  3. Carica il file della sequenza genetica.
  4. Ottiene la sequenza proteica dalla sequenza genetica.
  5. Carica e filtra le varianti genetiche per il gene specificato.
  6. Calcola e stampa alcune statistiche di base sulle varianti.
  7. Genera visualizzazioni e un report di riepilogo.
  8. Stampa un messaggio di completamento dell'analisi.
* **7. Entry point dello script**

A group of colorful lines

AI-generated content may be incorrect.

* Questa riga assicura che la funzione main venga eseguita solo se lo script viene eseguito direttamente, e non se viene importato come modulo in un altro script.