1. **dbsnp_vcf** and **dbsnp_vcf_idx**
    a. These contain common germline SNPs in vcf format
    b. These known variants are to ensure that common variants are not mistaken for sequencing errors and throw off recalibration
2. **known_indels** and **known_indels_idx**
    a. These contain common germline Indels in vcf format used in base recalibration
    b. These known variants are to ensure that common variants are not mistaken for sequencing errors and throw off recalibration
3. **mills_indels** and **mills_indels_idx**
    a. Another source of common germline indels in vcf format used in base recalibration
    b. These known variants are to ensure that common variants are not mistaken for sequencing errors and throw off recalibration
4. **gnomad** and **gnomad_idx**
    a. Provides germline variants from the Genome Aggregation Database and there allele frequencies needed for Mutect2 to calculate the likelihood of a variant being germline rather than somatic
    b. If we had a matched-normal sample, we would use that, but this is important for tumor-only mode
5. **filtered_vcf** and **filtered_vcf_idx**
    a. Common germline SNPs only (allele frequency > 5%) from the Exome Aggregation Consortium and used as filtering of variants
6. **pon** and **pon_idx**
    a. Panel of normals (PoN) from 1000 Genomes
    b. A PoN is used to filter out technical sources of variation and is recommended input for Mutect
    c. Because we are looking at technical bias, it is essential that the PoN is generated from a sequencing protocol as similar to the one used to generate your data
    d. Ideally this is from the same capture kit and at the same sequencing facility but if that is not available, the one I have provided is considered a good option as it is considered representative of the general population