Tempus Bioinformatics Technical Challenge

For this challenge, you are asked to prototype a variant annotation tool. We will provide you with a VCF file, and you will create a small software program to annotate each variant in the file.

Each variant must be annotated with the following pieces of information:

1. Type of variation (substitution, insertion, CNV, etc.) and their effect (missense, silent, intergenic, etc.). If there are multiple effects, annotate with the most deleterious possibility.
2. Depth of sequence coverage at the site of variation.
3. Number of reads supporting the variant.
4. Percentage of reads supporting the variant versus those supporting reference reads.
5. Allele frequency of variant from ExAC API (API documentation is available here: http://exac.hms.harvard.edu/).
6. Any additional annotations that you feel might be relevant.

Please upload all relevant code (written in whatever language you like) to a public repo on GitHub. Be sure to include the annotated variants in a csv/tsv file. Please also upload a README with the link to the repo via dropbox. Note that work will be assessed based on quality of code and documentation more-so than the annotation.

Feel free to reach out to bioinformatics-code-challenge@tempus.com with questions.