Evidences

Task completed using Perl 5, some requirements were not clear so I decided to add more functionalities to make sure that I achieve more than the minimum requirements requested.

1. Input and output file paths and names should be passed as script parameters rather than hard coded in the code.
   Example: `<script_name>.pl -i <Orignal CSV file> -o <Output CSV file>`

   To achieve this requirement I am using `Getopt::Long` to handle script parameters (`-i` for input file and `-o` for output file), which allows flexible input/output file paths.

- Running command using path to input file different from where the script is and writing the output file in another folder and receiving successful result.

- Running command using path to input file in the same folder of the script and writing the output file in the same folder, receiving successful result.

- Running command using a path that doesn't have a valid file and showing how to handle the error.

2. The output CSV file must contain the column headers which are missing from the original file, as well as the original data.

   To achieve this requirement, I ensure the output file contains headers (`column0`, `column1`, `column2`) which are added if missing from the original file.
   Since the headers name was not provided, I decided to use generic column names

```perl
my @header = map { "column$_" } (0 .. $max_columns - 1);
```

Input file with no Header

```
lpspalma@DESKTOP-TSVB697:/mnt/c/Users/luksp/OneDrive/Documentos/Lucas/Aulas.Tutoriais/perlReadWrite$ ls
 CSVProcessor.pm    Evidences.docx    Interview_task_1_Perl.docx    SampleCSVFile_11kb.csv    main.pl   '~$idences.docx'
 DataProcessor.pm   Evidences.pdf     README.md                     img
lpspalma@DESKTOP-TSVB697:/mnt/c/Users/luksp/OneDrive/Documentos/Lucas/Aulas.Tutoriais/perlReadWrite$ cat SampleCSVFile_11kb.csv
1,"Eldon Base for stackable storage shelf, platinum",Muhammed MacIntyre,3,-213.25,38.94,35,Nunavut,Storage & Organization,0.8
2,"1.7 Cubic Foot Compact ""Cube"" Office Refrigerators",Barry French,293,457.81,208.16,68.02,Nunavut,Appliances,0.58
3,"Cardinal Slant-D◆ Ring Binder, Heavy Gauge Vinyl",Barry French,293,46.71,8.69,2.99,Nunavut,Binders and Binder Accessories,0.39
4,R380,Clay Rozendal,483,1198.97,195.99,3.99,Nunavut,Telephones and Communication,0.58
5,Holmes HEPA Air Purifier,Carlos Soltero,515,30.94,21.78,5.94,Nunavut,Appliances,0.5
6,G.E. Longer-Life Indoor Recessed Floodlight Bulbs,Carlos Soltero,515,4.43,6.64,4.95,Nunavut,Office Furnishings,0.37
7,"Angle-D Binders with Locking Rings, Label Holders",Carl Jackson,613,-54.04,7.3,7.72,Nunavut,Binders and Binder Accessories,0.38
8,"SAFCO Mobile Desk Side File, Wire Frame",Carl Jackson,613,127.70,42.76,6.22,Nunavut,Storage & Organization,
9,"SAFCO Commercial Wire Shelving, Black",Monica Federle,643,-695.26,138.14,35,Nunavut,Storage & Organization,
10,Xerox 198,Dorothy Badders,678,-226.36,4.98,8.33,Nunavut,Paper,0.38
11,Xerox 1980,Neola Schneider,807,-166.85,4.28,6.18,Nunavut,Paper,0.4
12,Advantus Map Pennant Flags and Round Head Tacks,Neola Schneider,807,-14.33,3.95,2,Nunavut,Rubber Bands,0.53
13,Holmes HEPA Air Purifier,Carlos Daly,868,134.72,21.78,5.94,Nunavut,Appliances,0.5
14,"DS/HD IBM Formatted Diskettes, 200/Pack - Staples",Carlos Daly,868,114.46,47.98,3.61,Nunavut,Computer Peripherals,0.71
15,"Wilson Jones 1"" Hanging DublLock◆ Ring Binders",Claudia Miner,933,-4.72,5.28,2.99,Nunavut,Binders and Binder Accessories,0.37
16,Ultra Commercial Grade Dual Valve Door Closer,Neola Schneider,995,782.91,39.89,3.04,Nunavut,Office Furnishings,0.53
17,"#10-4 1/8"" x 9 1/2"" Premium Diagonal Seam Envelopes",Allen Rosenblatt,998,93.80,15.74,1.39,Nunavut,Envelopes,0.4
18,Hon 4-Shelf Metal Bookcases,Sylvia Foulston,1154,440.72,100.98,26.22,Nunavut,Bookcases,0.6
19,"Lesro Sheffield Collection Coffee Table, End Table, Center Table, Corner Table",Sylvia Foulston,1154,-481.04,71.37,69,Nunavut,Tables,0.68
20,g520,Jim Radford,1344,-11.68,65.99,5.26,Nunavut,Telephones and Communication,0.59
21,LX 788,Jim Radford,1344,313.58,155.99,8.99,Nunavut,Telephones and Communication,0.58
22,Avery 52,Carlos Soltero,1412,26.92,3.69,0.5,Nunavut,Labels,0.38
23,Plymouth Boxed Rubber Bands by Plymouth,Carlos Soltero,1412,-5.77,4.71,0.7,Nunavut,Rubber Bands,0.8
24,"GBC Pre-Punched Binding Paper, Plastic, White, 8-1/2"" x 11""",Carl Ludwig,1539,-172.88,15.99,13.18,Nunavut,Binders and Binder Accessories,0.37
25,"Maxell 3.5"" DS/HD IBM-Formatted Diskettes, 10/Pack",Carl Ludwig,1539,-144.55,4.89,4.93,Nunavut,Computer Peripherals,0.66
26,Newell 335,Don Miller,1540,5.76,2.88,0.7,Nunavut,Pens & Art Supplies,0.56
27,SANFORD Liquid Accent◆ Tank-Style Highlighters,Annie Cyprus,1702,4.90,2.84,0.93,Nunavut,Pens & Art Supplies,0.54
28,Canon PC940 Copier,Carl Ludwig,1761,-547.61,449.99,49,Nunavut,Copiers and Fax,0.38
29,"Tenex Personal Project File with Scoop Front Design, Black",Carlos Soltero,1792,-5.45,13.48,4.51,Nunavut,Storage & Organization,0.59
30,Col-Erase◆ Pencils with Erasers,Grant Carroll,2275,41.67,6.08,1.17,Nunavut,Pens & Art Supplies,0.56
31,"Imation 3.5"" DS/HD IBM Formatted Diskettes, 10/Pack",Don Miller,2277,-46.03,5.98,4.38,Nunavut,Computer Peripherals,0.75
32,"White Dual Perf Computer Printout Paper, 2700 Sheets, 1 Part, Heavyweight, 20 lbs., 14 7/8 x 11",Don Miller,2277,33.67,40.99,19.99,Nunavut,Paper,0.36
33,Self-Adhesive Address Labels for Typewriters by Universal,Alan Barnes,2532,140.01,7.31,0.49,Nunavut,Labels,0.38
34,Accessory37,Alan Barnes,2532,-78.96,20.99,2.5,Nunavut,Telephones and Communication,0.81
35,Fuji 5.2GB DVD-RAM,Jack Garza,2631,252.66,40.96,1.99,Nunavut,Computer Peripherals,0.55
36,Bevis Steel Folding Chairs,Julia West,2757,-1766.01,95.95,74.35,Nunavut,Chairs & Chairmats,0.57
37,Avery Binder Labels,Eugene Barchas,2791,-236.27,3.89,7.01,Nunavut,Binders and Binder Accessories,0.37
```

Output file with new header added

```
lpspalma@DESKTOP-TSVB697:/mnt/c/Users/luksp/OneDrive/Documentos/Lucas/Aulas.Tutoriais/perlReadWrite$ perl main.pl -i /mnt/c/Users/luksp/OneDrive/Documentos/
Lucas/Aulas.Tutoriais/SampleCSVFile_11kb.csv -o output.csv
Column column3:
  Sum: 503690
  Positive: 100 - 100.00%
  Negative: 0 - 0.00%
Column column4:
  Sum: 10452.15
  Positive: 49 - 49.00%
  Negative: 51 - 51.00%
Column column5:
  Sum: 7742.85
  Positive: 100 - 100.00%
  Negative: 0 - 0.00%
Column column6:
  Sum: 1350.1
  Positive: 100 - 100.00%
  Negative: 0 - 0.00%
CSV file has been processed successfully and written to 'output.csv'.
lpspalma@DESKTOP-TSVB697:/mnt/c/Users/luksp/OneDrive/Documentos/Lucas/Aulas.Tutoriais/perlReadWrite$ ls
 CSVProcessor.pm    Evidences.docx    Interview_task_1_Perl.docx    SampleCSVFile_11kb.csv    main.pl
 DataProcessor.pm   Evidences.pdf     README.md                     img                      output.csv   '~$idences.docx'
lpspalma@DESKTOP-TSVB697:/mnt/c/Users/luksp/OneDrive/Documentos/Lucas/Aulas.Tutoriais/perlReadWrite$ cat output.csv
column0,column1,column2,column3,column4,column5,column6,column7,column8,column9
1,"Eldon Base for stackable storage shelf, platinum","Muhammed MacIntyre",3,-213.25,38.94,35,Nunavut,"Storage & Organization",0.8
2,"1.7 Cubic Foot Compact ""Cube"" Office Refrigerators","Barry French",293,457.81,208.16,68.02,Nunavut,Appliances,0.58
3,"Cardinal Slant-D◆ Ring Binder, Heavy Gauge Vinyl","Barry French",293,46.71,8.69,2.99,Nunavut,"Binders and Binder Accessories",0.39
4,R380,"Clay Rozendal",483,1198.97,195.99,3.99,Nunavut,"Telephones and Communication",0.58
5,"Holmes HEPA Air Purifier","Carlos Soltero",515,30.94,21.78,5.94,Nunavut,Appliances,0.5
6,"G.E. Longer-Life Indoor Recessed Floodlight Bulbs","Carlos Soltero",515,4.43,6.64,4.95,Nunavut,"Office Furnishings",0.37
7,"Angle-D Binders with Locking Rings, Label Holders","Carl Jackson",613,-54.04,7.3,7.72,Nunavut,"Binders and Binder Accessories",0.38
8,"SAFCO Mobile Desk Side File, Wire Frame","Carl Jackson",613,127.70,42.76,6.22,Nunavut,"Storage & Organization",
9,"SAFCO Commercial Wire Shelving, Black","Monica Federle",643,-695.26,138.14,35,Nunavut,"Storage & Organization",
10,"Xerox 198","Dorothy Badders",678,-226.36,4.98,8.33,Nunavut,Paper,0.38
11,"Xerox 1980","Neola Schneider",807,-166.85,4.28,6.18,Nunavut,Paper,0.4
12,"Advantus Map Pennant Flags and Round Head Tacks","Neola Schneider",807,-14.33,3.95,2,Nunavut,"Rubber Bands",0.53
13,"Holmes HEPA Air Purifier","Carlos Daly",868,134.72,21.78,5.94,Nunavut,Appliances,0.5
```

3. The data in the original CSV must be parsed and stored inside a hash.
   The `read_csv_data` subroutine reads the CSV data into a hash (`%csv_hash`), storing the parsed data.

```perl
sub read_csv_data {
    my ($input_file) = @_;
    open my $input, '<', $input_file or die "Error: Could not open '$input_file': $!\n";
    my $csv = Text::CSV->new({ binary => 1, auto_diag => 1 });

    my @rows;
    my $max_columns = 0;
    while (my $row = $csv->getline($input)) {
        push @rows, $row;
        my $num_columns = scalar @$row;
        $max_columns = $num_columns if $num_columns > $max_columns;
    }
    close $input;

    my @header = map { "column$_" } (0 .. $max_columns - 1);

    my %csv_hash;
    my $row_index = 0;

    foreach my $row (@rows) {
        my %row_hash;
        @row_hash{@header} = (@$row, ('') x ($max_columns - @$row));
        $csv_hash{$row_index} = \%row_hash;
        $row_index++;
    }

    return (\%csv_hash, \@header);
}
```

4. The sum of the column 3 values must be printed out.
Although the script calculates sums, counts, and percentages for specified columns, it doesn't specifically print the sum of column 3 values. But the requirement was achieved since the script print the sum of the column 3 together with the other columns.

```
lpspalma@DESKTOP-TSVB697:/mnt/c/Users/luksp/OneDrive/Documentos/Lucas/Aulas.Tutoriais/perlReadWrite$ perl main.pl -i /mnt/c/Users/luksp/OneDrive/Documentos/
Lucas/Aulas.Tutoriais/SampleCSVFile_11kb.csv -o output.csv
Column column3:
  Sum: 503690
  Positive: 100 - 100.00%
  Negative: 0 - 0.00%
Column column4:
  Sum: 10452.15
  Positive: 49 - 49.00%
  Negative: 51 - 51.00%
Column column5:
  Sum: 7742.85
  Positive: 100 - 100.00%
  Negative: 0 - 0.00%
Column column6:
  Sum: 1350.1
  Positive: 100 - 100.00%
  Negative: 0 - 0.00%
CSV file has been processed successfully and written to 'output.csv'.
lpspalma@DESKTOP-TSVB697:/mnt/c/Users/luksp/OneDrive/Documentos/Lucas/Aulas.Tutoriais/perlReadWrite$ ls
```

In the result we are able to find the sum of the columns 3 to 6, columns that have numerical values, when is empty, the script read the field as 0.

5. The count of negative and positive column values should also display as well as the percentage of the total number of values.

Example: `Negative: 10 - 50%, Positive: 10 - 50%`

The `calculate_data` subroutine calculates counts and percentages for positive and negative values in specified columns, which meets this requirement.

```
lpspalma@DESKTOP-TSVB697:/mnt/c/Users/luksp/OneDrive/Documentos/Lucas/Aulas.Tutoriais/perlReadWrite$ perl main.pl -i /mnt/c/Users/luksp/OneDrive/Documentos/
Lucas/Aulas.Tutoriais/SampleCSVFile_11kb.csv -o output.csv
Column_column3:
  Sum: 503690
  Positive: 100 - 100.00%
  Negative: 0 - 0.00%
Column_column4:
  Sum: 10452.15
  Positive: 49 - 49.00%
  Negative: 51 - 51.00%
Column_column5:
  Sum: 7742.85
  Positive: 100 - 100.00%
  Negative: 0 - 0.00%
Column_column6:
  Sum: 1350.1
  Positive: 100 - 100.00%
  Negative: 0 - 0.00%
CSV file has been processed successfully and written to 'output.csv'.
```

Error Handling and Clear Messages: I've implemented error handling for file operations (die statements with clear error messages) and provided appropriate error messages. The error messages indicate file open failures (Could not open '<filename>' for writing: <error>).

Documentation and Structure: I added comments and structured subroutines (main, read_csv_data, calculate_data, write_output, etc.), which enhances readability and maintainability. Also, the Readme doc was added in the project detailing the execution, function details and test strategy.

Test Plan: Unit tests (Test::More) that cover positive, negative, and edge scenarios for both most part of the functions. These tests validate functionality and handle error cases effectively.

```
lpspalma@DESKTOP-TSVB697:/mnt/c/Users/luksp/OneDrive/Documentos/Lucas/Aulas.Tutoriais/perlReadWrite$ prove t/*.t
t/CSVProcessorTest.t ... ok
t/DataProcessorTest.t .. ok
All tests successful.
Files=2, Tests=22,  0 wallclock secs ( 0.06 usr  0.00 sys +  0.23 cusr  0.02 csys =  0.31 CPU)
Result: PASS
lpspalma@DESKTOP-TSVB697:/mnt/c/Users/luksp/OneDrive/Documentos/Lucas/Aulas.Tutoriais/perlReadWrite$
```

```
lpspalma@DESKTOP-TSVB697:/mnt/c/Users/luksp/OneDrive/Documentos/Lucas/Aulas.Tutoriais/perlReadWrite$ prove t/*.t
t/CSVProcessorTest.t ... ok
t/DataProcessorTest.t .. ok
All tests successful.
Files=2, Tests=22,  0 wallclock secs ( 0.06 usr  0.00 sys +  0.23 cusr  0.02 csys =  0.31 CPU)
Result: PASS
lpspalma@DESKTOP-TSVB697:/mnt/c/Users/luksp/OneDrive/Documentos/Lucas/Aulas.Tutoriais/perlReadWrite$ perl t/CSVProcessorTest.t
1..12
ok 1 - Header should match
ok 2 - First row should match
ok 3 - Second row should match
ok 4 - Header should be empty
ok 5 - CSV hash should be empty
ok 6 - Header should match
ok 7 - First row should match with padding
ok 8 - Second row should match
ok 9 - Third row should match with padding
ok 10 - Output file should match expected content
ok 11 - Correct error message for permission denied
ok 12 - Correct error message for non-existent directory
lpspalma@DESKTOP-TSVB697:/mnt/c/Users/luksp/OneDrive/Documentos/Lucas/Aulas.Tutoriais/perlReadWrite$ perl t/DTest.t
Can't open perl script "t/DTest.t": No such file or directory
lpspalma@DESKTOP-TSVB697:/mnt/c/Users/luksp/OneDrive/Documentos/Lucas/Aulas.Tutoriais/perlReadWrite$ perl t/DataProcessorTest.t
ok 1 - Sum calculation for column1 is correct
ok 2 - Sum calculation for column2 is correct
ok 3 - Sum calculation for column3 is correct
ok 4 - Count positive values for column1 is correct
ok 5 - Count negative values for column1 is correct
ok 6 - Total count for column1 is correct
ok 7 - Percentage calculation for positive values in column1 is correct
ok 8 - Percentage calculation for negative values in column1 is correct
ok 9 - Empty data returns zero sums
ok 10 - No negative values count is correct
1..10
lpspalma@DESKTOP-TSVB697:/mnt/c/Users/luksp/OneDrive/Documentos/Lucas/Aulas.Tutoriais/perlReadWrite$
```