

A Specification: The Tax Code

A.1 Background

You live in the country of Avalon. Spring is here and the smell of taxes fills the air. Like the United States, citizens of Avalon pay taxes both to their federal government and to the state in which they reside.

Avalon has recently modernized its tax system. There is a new federal tax code (section A.2) and a new tax code for each state (section A.3). However, the official tax calculator is proprietary and not publicly available. You would like to implement your own tax calculator so that you do not have to compute your taxes by hand.

Your task is to implement a tax computation engine that reads an input file describing taxpayers' financial information and outputs the total tax that each one owes under the 2026 Avalon tax code (federal + state).

A.2 Avalon Federal Tax Code

Taxes are taken on the adjusted income, which is derived from the gross income and a series of deductions.

A.2.1 Gross Income. Gross income is defined as the sum of all taxable income sources before deductions or credits are applied. Gross income consists of the following components:

- **Wage Income (W2).** The total annual wage income reported by the employer. This value is provided directly in the input and is treated as fully taxable.
- **Capital Gains and Losses.** Net income resulting from the sale of financial assets. Capital gain for a transaction is defined as:

$$\text{Gain} = \text{Sale Price} - \text{Purchase Price}.$$

Losses occur when this quantity is negative.

Gross income is computed as:

$$\text{Gross Income} = \text{Wage Income} + \text{Net Capital Gain}.$$

Net capital gain combines the gains and losses from the reporting year.

A.2.2 Capital Gains Computation. Taxpayers may have invested in one of the three publicly traded companies in Avalon: Ananya's Wool Whimsy, Miko's Coal Collaborative, and Esteban's Timberfell. If a taxpayer sells any portion of an investment during the reporting year, the gain or loss from that sale (delta from the purchase price) is taxable.

Realized Gain or Loss. For each sale of an asset, the realized gain (or loss) is defined as:

$$\text{Gain} = q \cdot (p_s - p_b),$$

where:

- q is the quantity of shares sold,
- p_s is the sale price per share,
- p_b is the purchase price per share.

Cost Basis Matching Rule. If shares in one of the companies were purchased at different times of the year, the purchase price may have varied. To calculate gains, sales must be matched to prior purchases using the **FIFO** (first-in, first-out) rule. That is, shares sold are matched against the earliest purchased shares that have not yet been sold.

Net Capital Gain. The taxpayer's **Net Capital Gain** is the sum of realized gains and losses across all assets and all sales during the reporting year:

$$\text{Net Capital Gain} = \sum_{\text{all realized transactions}} \text{Gain.}$$

All transactions are conducted to three decimal places. Results should be rounded to the nearest Avalon dollar.

A.2.3 Adjustments and Deductions. After computing gross income, the taxpayer may be eligible for certain adjustments and deductions. These reduce income before the application of tax brackets and surcharge calculations.

Deductions are applied in the order specified below. All deductions reduce gross income to produce **Taxable Income**.

Any excess deduction beyond gross income is discarded and does not carry forward.

A taxpayer may either take the Standard Deduction or itemize deductions.

Standard Deduction. The Standard Deduction is a fixed dollar amount: \$10,000. If elected, it reduces gross income directly. If the Standard Deduction is taken, no itemized deductions may be applied.

Itemized Deductions. Instead of taking the Standard Deduction, a taxpayer may elect to itemize deductions. If itemizing, the taxpayer may claim:

- Charitable Donation Deduction,
- Child Tax Deduction.

Itemized deductions reduce gross income before bracket computation.

Charitable Donation Deduction. The taxpayer may report a list of charitable donations made during the reporting year. Each donation is specified as a monetary amount in dollars.

The Charitable Donation Deduction is equal to the sum of all reported donations.

$$\text{Charitable Deduction} = \sum_{i=1}^n d_i$$

where d_i are the individual donation amounts.

Child Tax Deduction. A taxpayer may claim a deduction based on the number of qualifying children. The deduction percentage increases with each additional child, up to a maximum of 10 children.

The applicable percentage schedule is:

- 1% for the first child,
- 2% for the second child,
- 3% for the third child,
- ...
- up to 10% for the tenth child.

The deduction is computed as a percentage of taxable income before the child deduction is applied.

A.2.4 Taxable Income. Formal definition:

$$\text{Taxable Income} = \text{Gross Income} - \text{Allowable Deductions}.$$

The taxpayer is obligated to select whichever option (Standard or Itemized) results in the lower total tax owed.

A.2.5 Progressive Tax Brackets. National income tax is applied to the Taxable income according to the following marginal bracket structure:

Taxable Income Range	Marginal Rate
$0 \leq I \leq 100,000$	5%
$100,000 < I \leq 200,000$	10%
$200,000 < I \leq 300,000$	15%
$I > 300,000$	20%

Policy Note: This bracket structure is intended to incentivize citizens to make more money by maintaining relatively modest increases in marginal tax rates at higher income levels.

A.2.6 High-Income Surcharge. In addition to the national income tax computed under the progressive bracket structure, a High-Income Surcharge may apply.

This surcharge is computed independently of deductions and bracket calculations. It is applied directly to **Gross Income**.

Five-Year Income History. Each taxpayer provides reported gross income values for the previous five tax years:

$$Y_{t-1}, Y_{t-2}, Y_{t-3}, Y_{t-4}, Y_{t-5}.$$

These values are expressed in dollars and may include fractional cents.

Exponentially Weighted Moving Average (EWMA). The five-year average income is computed using an exponentially weighted moving average with smoothing factor $\alpha = 0.6$.

The EWMA is defined recursively as:

$$E_1 = Y_{t-5}$$

$$E_k = \alpha Y_{t-(6-k)} + (1 - \alpha)E_{k-1} \quad \text{for } k = 2, 3, 4, 5.$$

The final EWMA value is E_5 .

Threshold for Surcharge. If

$$E_5 > 1,000,000,$$

the taxpayer is subject to the High-Income Surcharge.

Surcharge Computation. If applicable, the surcharge is computed as:

$$\text{Surcharge} = 0.02 \times \text{Gross Income}.$$

If the EWMA threshold is not exceeded, the surcharge is zero.

Total National Tax. The total national tax is:

$$\text{National Tax} = \text{Bracket Tax (on Taxable Income)} + \text{Surcharge}.$$

A.3 Avalon State Tax Code

Each taxpayer resides in exactly one state. They cannot be employed in any other state, nor leave the state at any time, thus simplifying the tax code (although perhaps not their lives).

State tax is computed independently of national tax and added to the total tax owed. The two states of Avalon are:

- California
- Texas

A.3.1 California State Tax. California computes state tax based on **Gross Income**, but treats income components differently.

State tax is defined as:

$$\text{CA Tax} = 0.04 \cdot \text{Wage Income} + 0.06 \cdot \max(0, \text{Net Capital Gain})$$

Notes:

- Wage income is taxed at a flat rate of 4%.
- Positive net capital gain is taxed at 6%.
- If net capital gain is negative, it contributes 0 to the capital-gains component of state tax.
- No deductions (standard, charitable, or child) apply to California state tax.
- If the federal High-Income Surcharge applies to a California resident, a 5% surcharge is also applied at the state level.

A.3.2 Texas State Tax. Texas computes state tax based on **Taxable Income** (after deductions).

Texas uses the following marginal bracket structure:

Taxable Income Range	Marginal Rate
$0 \leq I \leq 90,000$	3%
$90,000 < I \leq 200,000$	5%
$I > 200,000$	7%

Tax is computed marginally.

Notes:

- Texas state tax applies to Taxable Income as defined in the national tax section.
- If their federal deductions exceed 15,000, then citizens shall apply a 1% deduction to their computed state tax.
- In Texas, the High-Income Surcharge does not affect state tax computation.

A.3.3 Total Tax Liability. Total tax owed is:

$$\text{Total Tax} = \text{National Tax} + \text{State Tax}.$$

A.4 Numerical Precision

All taxes shall be paid in whole dollars.

A.5 Input and Output Format Specification

A.5.1 Input Format (NDJSON). The input file shall be in **NDJSON** (newline-delimited JSON) format. Each line of the file is a valid JSON object representing a single household.

Each household object contains four conceptual components:

- Household data,
- Earned income history,
- Investment data,
- Charitable giving.

Household Data.

- `taxpayer_id` (string)
- `state` (string: "California" or "Texas")
- `w2_income` (number)
- `num_children` (integer)

Earned Income History.

- `prior_five_years_income` (array of exactly five numbers)

The array is ordered from oldest to most recent year:

$$[Y_{t-5}, Y_{t-4}, Y_{t-3}, Y_{t-2}, Y_{t-1}]$$

Investment Data. Investment activity is represented using two arrays:

- `purchases`: array of objects
- `sales`: array of objects

Each purchase object contains:

- `asset_id` (string)
- `date` (string, ISO-8601 format: YYYY-MM-DD)
- `quantity` (number)
- `unit_price` (number)

Each sale object contains:

- `asset_id` (string)
- `date` (string, ISO-8601 format: YYYY-MM-DD)
- `quantity` (number)
- `unit_price` (number)

Charitable Giving. Charitable giving is represented as:

- `charitable_donations`: array of numbers

Each element of the array represents the monetary value (in dollars and cents) of a single charitable donation made during the reporting year.

A.5.2 Output Format (NDJSON). The output file shall also be in **NDJSON** (newline-delimited JSON) format. Each line of the output corresponds one-to-one with a household object in the input file.

For every input household object, exactly one output JSON object must be produced.

Each output object contains the following fields:

- `taxpayer_id` (string),
- `federal_tax` (number),
- `state_tax` (number).

The order of output taxes due must match the order of households in the input file.

A.5.3 CLI. Each implementation should be `chmod +x`'d so the runner can invoke it, and must accept the following command-line arguments:

- `-inputFile HOUSEHOLDS.NDJSON` : specifies the path to an NDJSON input file (one JSON object per line).
- `-outputFile TAXES.NDJSON` : specifies the path where the implementation must write its NDJSON output.

The program must read from the given input file and write all results to the given output file. No other input/output behavior should be required.

B Sample Prompts

This appendix provides sample prompts you may use when interacting with an LLM. You are not required to use these prompts verbatim. However, you should keep the specification fixed and avoid importing external tax rules beyond what is stated in section A.

B.1 Prompt 1: Increasing Design Diversity

Prompt: Hi ChatGPT, I'm interested in developing diverse software implementations of a spec. I am thinking about ways to prompt you that will increase the variation in the implementations. I am wondering about programming language and library selection (vs. custom implementations of components). Are there other things I should think about? Please give concrete strategies and examples, and explain how each strategy might affect the diversity of failure modes.

B.2 Prompt 2: Starting from a System Design

Prompt: The full spec is available as a file in your context. Can we start with a system design? I want the design to have an NDJSON parser, an analysis engine, and an output writer. I would also like to use a database to store intermediate computations, so that the tax computation pipeline is explicit and debuggable. Please propose a modular architecture and data model first, before writing any code. Then, after I confirm, we will implement the modules one at a time.

References

- [1] Mark AI. 2025. Transformers Agents Tutorial: Building AI Assistants with Tool Use. <https://markaicode.com/transformers-agents-tutorial-ai-assistants-tool-use/>
- [2] Microsoft Corporation. 2026. Example Skills. https://code.visualstudio.com/docs/copilot/customization/agent-skills#_example-skills
- [3] Claude Inc. 2026. Agent SDK: Overview. <https://platform.claude.com/docs/en/agent-sdk/overview>
- [4] Firstname LASTNAME and Firstname LASTNAME. YYYY. TITLE OF SOURCE. <https://example.com/> Accessed: YYYY-MM-DD.