

Personal Income Tax Filers: *Summary datasets* -- beginning Tax Year 2015

(8/18) **Overview for Four (4) datasets:**

- 1) *Major Items by Liability Status and Place of Residence***
- 2) *Major Items and Income & Deduction Components by Place of Residence and Broad Income Range***
- 3) *Statewide Major Items and Income & Deduction Components by Liability Status and Detail Income Range***
- 4) *County-level Major Items and Income & Deduction Components by Wide Income Range***

General Description

Beginning with tax year 2015, the Department of Taxation and Finance (hereafter “the Department”) began producing a new annual population data study file to provide more comprehensive statistical information on New York State personal income tax returns. The data are from full-year resident, nonresident, and part-year resident returns filed between January 1 and December 31 of the year after the start of the liability period (hereafter referred to as the “processing year”). The four datasets display major income tax components by tax year. This includes the distribution of New York adjusted gross income and tax liability by county or place of residence, as well as the value of deductions, exemptions, taxable income and tax before credits by size of income. In addition, three of the four datasets include all the components of income, the components of deductions, and the addition/subtraction modifications.

Data Collection Methodology

The datasets include summarized data from the tax year’s population of all returns received during the processing year. The personal income tax return information in these datasets represents the compilation of records from the Department’s processing systems at a point in time. The Department’s Office of Tax Policy Analysis compiles and reviews the data and reports the statistics.

For the purposes of these datasets, all returns are categorized into one of three residency statuses – full-year NYS resident, full-year nonresident, or part-year resident – and one of two tax liability statuses explained later. Residency status is determined by both the type of form filed and by self-reported indicators. Generally, full year-resident filers are individuals who live in New York during the entire year and file the tax form IT-201 called a resident tax return, in contrast to full-year nonresidents and part-year residents, who file the IT-203. However, some full-year residents have a permanent home (i.e. a domicile) in New York, but can live in another state or country for most of the year. Other residents have a permanent home outside the state, but maintain living quarters in New York State and are present in New York for more than half of the tax year. Likewise, discerning from a return the permanent home of a nonresident or part-year resident can result in some generally categorized places, such as “Other Places,” as explained in the data dictionaries. The county or place of residence for filers is as of December 31 of the applicable tax year.

Along with residency status, all returns are categorized into one of two tax liability statuses – either taxable or non-taxable. The tax liability reflects the tax after credits; the value can be negative due to the impact of refundable tax credits. Returns that have a positive tax liability after having been processed by the Department

are defined as taxable returns. In contrast, returns that do not have a positive tax liability, that is, returns with a zero or negative tax liability, are defined as non-taxable returns.

Statistical and Analytic Issues

The datasets are updated annually on a routine basis for the next tax year once returns are filed, processed, and verified. Due to the nature of personal income tax return filing and subsequent processing and data collection, the reporting of the data is typically lagged by two years.

We recommend no comparison be done with any other statistical reports or datasets created by the Department for tax years 2014 and prior. For example, in contrast to the *Analysis of Personal Income Tax Returns* reports published in tax years prior to 2015, which were produced from a sample file of returns for primarily full-year resident taxable returns, these datasets are created using a population study file of more than 10 million annual returns. Thus, in addition to full year resident and taxable returns, the datasets include nonresident and part-year resident returns, as well as taxable and non-taxable returns for all residency types. One significant difference and benefit to moving from a sampled file to the population file is the removal of sampling or weighting bias that allows us to report component data at the county level. In summary, the inclusion of all residency types for both taxable and non-taxable returns, plus all component data, allows for a more complete representation of the population and its analyses than in prior reports and datasets.

Limitations of Data Use

The four datasets present various and distinct distributions of the data from the population of New York State personal income tax filers. However, due to tax secrecy provisions, certain data cannot be disclosed; in those cases, fields are blanked out. In some instances, only a few fields of one summarized record are necessarily omitted (as Noted with d/). In instances where the nondisclosure was extensive or across numerous records and variables, entire sections are blanked out (as Noted with a numerical footnote) and only the totals are provided. The information reflects personal income tax return and amended return data received during the processing year. Therefore, the information contained in these datasets does not reflect final values for all returns for any tax year.

The liability amounts contained in these datasets do not reflect actual tax collections. Total tax collections are reported on a State fiscal basis (April 1 through March 31) and include receipts of various types of payments made to satisfy taxpayers' tax liability – including withholding, estimated payments, final payments, refunds, audit payments, and payments from delinquent accounts. Tax collections are reported in the Department's [Fiscal Year Tax Collections reports](#), and not in these datasets.