

Información para dar cumplimiento al Artículo Décimo Noveno Transitorio de la LIF 2017

En cumplimiento con lo establecido en el Artículo Décimo Noveno Transitorio de la Ley de Ingresos de la Federación para 2017¹, el cual establece lo siguiente:

“Con el propósito de coadyuvar a la transparencia y al análisis de la política fiscal, así como de facilitar el estudio de la distribución del ingreso entre la población y del efecto de la política tributaria sobre dicha distribución, en el ejercicio fiscal 2017 la Secretaría de Hacienda y Crédito Público, a través del Servicio de Administración Tributaria, deberá hacer pública información estadística anónima de las declaraciones anuales del impuesto sobre la renta de los contribuyentes personas físicas y morales. En ningún caso, la información pública podrá contener datos sobre el nombre, denominación o razón social y clave del registro federal de contribuyentes o información que permita la identificación del contribuyente.

El Servicio de Administración Tributaria determinará las características, fuentes, metodología y periodo de la información estadística a publicar, misma que dará a conocer en su página de Internet. La información estadística será actualizada cada dos años.”

En este sentido, el presente documento contiene las características generales de la información estadística a que se refiere dicho ordenamiento.

La información fiscal se sujetó a un proceso de anonimización² para ponerse a disposición del público en general considerando los siguientes aspectos:

- A) La información se presenta a nivel de micro dato, es decir, contiene la información abierta de cada uno de los campos seleccionados de la declaración de cada contribuyente.
- B) Confidencialidad. No se adjunta información de carácter cualitativo, es decir, no se publica información sobre cualidades del contribuyente (entidad, sector, domicilio, etc.), sólo se incluye un identificador único, por contribuyentes que es el mismo en todo el periodo.
- C) Se respeta el principio de autodeterminación. La información anonimizada respeta los datos que plasmó el contribuyente en su declaración, aún y cuando sean inconsistentes³. Sin embargo, en lo general, la información refleja la normatividad aplicable.

¹ El texto completo se encuentra en http://www.dof.gob.mx/nota_detalle.php?codigo=5460968&fecha=15/11/2016

² Se entiende como anonimización, el proceso por el cual deja de ser posible establecer por medios razonables el nexo entre un dato y el sujeto al que se refiere.

³ La autodeterminación dispone que corresponde a los contribuyentes la determinación del pago de los impuestos a su cargo, salvo disposición expresa en contrario. Consecuentemente, en el causante recae la responsabilidad de determinar en cantidad líquida las contribuciones a enterar, mediante la realización de las operaciones matemáticas encaminadas a fijar su importe exacto a través de la aplicación de las tasas tributarias establecidas en la Legislación tributaria. Dicha autodeterminación de las contribuciones parte de un principio de buena fe, el cual le permite declarar voluntariamente el monto de sus obligaciones fiscales.

Información para dar cumplimiento al Artículo Décimo Noveno Transitorio de la LIF 2017

Fuentes

La información a que se refiere el artículo Décimo Noveno transitorio de la LIF 2017 proviene de las declaraciones fiscales y se publicará en formato de bases de datos en dos fases:

- Fase 1 Declaración anual de Personas Morales y Personas Físicas
 - Régimen General
 - Obligados y no obligados
- Fase 2. Declaración Informativa Múltiple de sueldos y salarios.
 - Asalariados

Período

La información de cada fase será publicada de manera escalonada durante 2017.

FASE 1: Declaración anual de Personas Morales Régimen General

En esta etapa, se publicarán datos de las declaraciones anuales de los contribuyentes que tributan bajo el Régimen General, de los ejercicios 2010 al 2015. Se seleccionó el conjunto de 36 variables que contemplan la integración del ISR.

#	Variable	#	Variable
1	TOTAL DE INGRESOS ACUMULABLES	19	TOTAL DE ESTIMULOS
2	TOTAL DE DEDUCCIONES AUTORIZADAS Y DEDUCCIÓN INMEDIATA DE INVERSIONES	20	PAGOS PROVISIONALES EFECTUADOS ENTREGADOS A LA CONTROLADORA
3	UTILIDAD O PERDIDA FISCAL ANTES DE PTU	21	PAGOS PROVISIONALES EFECTUADOS ENTERADOS A LA FEDERACIÓN
4	PTU PAGADA EN EL EJERCICIO	22	IMPUESTO RETENIDO AL CONTRIBUYENTE
5	UTILIDAD FISCAL DEL EJERCICIO	23	IMPUESTO ACREDITARLE PAGADO EN EL EXTRANJERO
6	PERDIDA FISCAL DEL EJERCICIO	24	IMPUESTO ACREDITARLE POR DIVIDENDOS O UTILIDADES DISTRIBUIDOS
7	PERDIDAS FISCALES DE EJERCICIOS ANTERIORES QUE SE APLICAN EN EL EJERCICIO	25	CRÉDITO FISCAL IETU POR DEDUCCIONES MAYORES A LOS INGRESOS
8	DEDUCCIÓN ADICIONAL DEL FOMENTO AL PRIMER EMPLEO	26	IMPUESTO CORRESPONDIENTE A LA CONSOLIDACIÓN FISCAL A CARGO
9	RESULTADO FISCAL	27	IMPUESTO CORRESPONDIENTE A LA CONSOLIDACION FISCAL ENTREGADO (en exceso)
10	IMPUESTO SOBRE LA RENTA DEL EJERCICIO	28	OTRAS CANTIDADES A CARGO
11	REDUCCIONES PARA MAQUILADORAS	29	OTRAS CANTIDADES A FAVOR
12	CREDITO FISCAL POR DEDUCCION INMEDIATA PARA MAQUILADORAS	30	DIFERENCIA A CARGO
13	OTRAS REDUCCIONES DEL ISR	31	DIFERENCIA A FAVOR
14	IMPUESTO CAUSADO EN EL EJERCICIO	32	IMPUESTO ACREDITABLE POR DEPÓSITOS EN EFECTIVO DEL EJERCICIO
15	ESTIMULO POR PROYECTOS EN INVESTIGACIÓN Y DESARROLLO TECNOLÓGICO	33	IMPUESTO A LA VENTA DE BIENES Y SERVICIOS Suntuarios ACREDITARLE
16	ESTIMULO POR PROYECTOS DE INVERSIÓN EN LA PRODUCCIÓN CINEMATOGRAFICA NACIONAL	34	ISR A CARGO DEL EJERCICIO
17	ESTIMULO A PROYECTOS DE INVERSIÓN EN LA PRODUCCIÓN TEATRAL NACIONAL	35	ISR PAGADO EN EXCESO APLICADO CONTRA EL IMPAC
18	OTROS ESTIMULOS	36	ISR A FAVOR DEL EJERCICIO

Información para dar cumplimiento al Artículo Décimo Noveno Transitorio de la LIF 2017

FASE 1: Declaración anual de Personas Físicas Obligados y No Obligados

En esta etapa, se publicarán datos de las declaraciones anuales de los contribuyentes clasificados como Personas Físicas, de los ejercicios 2010 al 2015. Se seleccionó el conjunto de 36 variables que contemplan la integración del ISR. Las variables 32 al 43 son las aperturas del “ingreso o utilidades acumulables”.

#	Variable	#	Variable
1	INGRESOS O UTILIDADES ACUMULABLES	23	OTRAS CANTIDADES A CARGO DEL CONTRIBUYENTE
2	PERDIDAS	24	OTRAS CANTIDADES A FAVOR DEL CONTRIBUYENTE
3	TOTAL DE INGRESOS ACUMULABLES	25	IMPUESTO A CARGO (1A DIFERENCIA)
4	DEDUCCIONES PERSONALES	26	ISR PAGADO EN EXCESO
5	BASE GRAVABLE	27	IMPUESTO ACREDITABLE POR DEPOSITOS EN EFECTIVO DEL EJERCICIO
6	ISR CONFORME A TARIFA ANUAL	28	ISR A CARGO DEL EJERCICIO
7	SUBSIDIO ACREDITABLE	29	IDE PENDIENTE DE APLICAR DEL EJERCICIO
8	SUBSIDIO PARA EL EMPLEO	30	ISR PAGADO EN EXCESO APLICADO CONTRA EL IMPAC
9	IMPUESTO SOBRE INGRESOS ACUMULABLES	31	ISR A FAVOR DEL EJERCICIO
10	IMPUESTO SOBRE INGRESOS NO ACUMULABLES	32	INGRESOS ACUMULABLES- Sección Ingresos por Salarios
11	REDUCCIONES DE ISR	33	INGRESOS ACUMULABLES- Sección Arrendamiento
12	IMPUESTO SOBRE LA RENTA CAUSADO	34	GANANCIA ACUMULABLE- Sección Enajenación de Bienes
13	ESTIMULOS FISCALES APLICADOS DURANTE EL EJERCICIO	35	INTERÉS REAL PROVENIENTES DEL SISTEMA FINANCIERO- Sección Intereses
14	ESTIMULO POR PROYECTOS DE INVESTIGACION Y DESARROLLO TECNOLÓGICO	36	INTERÉS REAL NO PROVENIENTES DEL SISTEMA FINANCIERO- Sección Intereses
15	ESTIMULO POR PROYECTOS DE INVERSION EN LA PRODUCCION CINEMATOGRAFICA NACIONAL	37	INGRESOS ACUMULABLES POR DIVIDENDOS O UTILIDADES- Sección Dividendos o Utilidades
16	ESTIMULO POR PROYECTOS DE INVERSION EN LA PRODUCCION TEATRAL NACIONAL	38	INGRESOS ACUMULABLES - Sección De los demás ingresos
17	PAGOS EFECTUADOS A LA ENTIDAD FEDERATIVA	39	UTILIDAD GRAVABLE ACUMULABLE - Sección Actividades Profesionales (Sin actividad empresarial)
18	PAGOS PROVISIONALES EFECTUADOS A LA FEDERACION	40	INGRESOS ACUMULABLES" Sección Adquisición de Bienes
19	IMPUESTO RETENIDO AL CONTRIBUYENTE O PAGADO POR TERCEROS	41	INGRESOS ACUMULABLES POR PREMIOS- Sección Premios
20	IMPUESTO ACREDITABLE PAGADO EN EL EXTRANJERO	42	UTILIDAD GRAVABLE ACUMULABLE- Sección Régimen Intermedio de Actividades Empresariales (ISR)
21	CREDITO FISCAL IETU POR DEDUCCIONES MAYORES A LOS INGRESOS	43	UTILIDAD GRAVABLE ACUMULABLE- Sección Resultados
22	IMPUESTO SOBRE EL INTERES REAL POR RETIROS PARCIALES		

Información para dar cumplimiento al Artículo Décimo Noveno Transitorio de la LIF 2017

Metodología; Anonimización a través de ruidos.

Para mantener la confidencialidad del contribuyente, la información a publicar será anonimizada, es decir, a partir de la información real, se genera un nuevo dato, el cual respeta propiedades estadísticas de los datos en su conjunto. Para ello se utilizaron técnicas que enmascaran los datos originales a través de la aplicación de choques aleatorios conocidos como ruidos, con el fin de eliminar la relación entre individuos y datos.

El valor del ruido es elegido de una distribución normal con media cero y una desviación estándar, de tal forma que este valor no distorsione estadísticos como la media, desviación estándar y coeficientes de correlación entre variables. Dichos ruidos son calculados a partir de las bases originales, la expresión se puede expresar como:

$$Y = X + \varepsilon$$

Donde Y es el valor anonimizado, X es el valor original y ε es la variable aleatoria (ruido) con una distribución $\varepsilon \sim N(0, \sigma^2)$.

Para generar la variable anonimizada, se le suma el ruido obtenido a la variable original, siempre y cuando la variable original sea mayor a "0", si esta condición no se cumple, se asume que la variable anonimizada es "0".

Por otro lado, se respeta el sentido de las variables originales (por ejemplo, no pueden existir ingresos negativos), de esta forma, si la variable original no contempla negativos, el valor anonimizado tampoco los contemplará.

Asimismo, si dentro de la declaración el contribuyente no presentó dato (esto expresado como valor nulo) la anonimización respetará dicho valor, por ende, el valor anonimizado será nulo.

Como parte del proceso de anonimización, los datos individuales pueden perder la relación con los cálculos fiscales, no obstante, en términos agregados se conservan las propiedades estadísticas del conjunto de datos.

Limpieza y estandarización

Para cada ejercicio fiscal, se toma la última declaración presentada por el contribuyente al último corte de la extracción de datos. Derivado de los análisis realizados en la creación del modelo y con la finalidad de disminuir las varianzas entre las variables y disminuir las discrepancias en las propiedades estadísticas de las fuentes, fue necesario eliminar registros que producían dichos incrementos.

Para ello con base en el "Total de ingresos acumulables" se estableció un criterio de eliminación de registros atípicos, mediante el cálculo de la media (μ_{Ti}) y la desviación estándar (DST), eliminando aquellas observaciones que estuvieran por arriba de la media más tres desviaciones estándar ($\mu_{Ti} + 3 * DST$).

Información para dar cumplimiento al Artículo Décimo Noveno Transitorio de la LIF 2017

Preguntas frecuentes

¿Cuál es la diferencia entre valores nulos y cero?

R: El valor nulo expresa ausencia de valor (ya sea porque el contribuyente no llenó el campo o el aplicativo no realizó el cálculo por ausencia de valores, estos valores están expresados con "."). En cambio, el valor en "0" significa que el contribuyente declaró en "0" la variable o que el resultado automático del cálculo fue "0". "

¿Por qué existen igualdades lógicas que no se cumplen como Ingresos-Deducciones=Utilidad antes PTU?

R: Se anonimizaron los valores declarados por los contribuyentes, omitiendo sus relaciones, por ende, no se construyen variables a partir de los datos anonimizados. Sin embargo, este criterio no altera las relaciones funcionales y estadísticas de los agregados, aunque sí las relaciones a nivel contribuyente.

Por otro lado, se respeta la "autodeterminación fiscal" por lo que de origen los datos pueden ser incongruentes.

¿El identificador del RFC es el mismo para todos los años?

R: Sí, los identificadores del RFC representan al mismo individuo para todos los ejercicios presentados.

¿Por qué si sumo todos los impuestos de un ejercicio, el resultado no es parecido al dato publicado por el SAT o la SHCP?

R: El proceso de anonimización genera que los datos tengan relación con los datos originales, guardando cierta proporcionalidad con los datos originales, sin embargo, el dato original no es el mismo.

Por otro lado, durante los procesos de limpieza y estandarización, se pierden datos originales. Así mismo, las tablas de declaraciones son dinámicas, por lo que la información es una fotografía de determinado día y la variación de cifras es normal por los ajustes u operaciones posteriores que realice el contribuyente.

Cabe hacer mención que, para estos ejercicios, se toman la última declaración presentada por los contribuyentes.