



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Horas trabajadas: un análisis integral

PROYECTO I: Ciencia de Datos. Grupo 8

Curso 2021/2022



Javier Luque, Daniel Garijo, Pablo Parrilla, José Valero y Júlia Vericat

Presentación

Este proyecto ha sido realizado por alumnos de la ETSINF (Escuela Técnica Superior de Ingeniería Informática) de la Universitat Politècnica de València. Además, se ha realizado en la asignatura de Proyecto I, del segundo cuatrimestre del primer curso del Grado de “Ciencia de Datos”. A lo largo del cuatrimestre, nuestro equipo ha estado trabajando en este estudio acerca de las horas trabajadas y su posible situación en el futuro. Este trabajo ha dado como resultado esta memoria, así como una página web y una serie de análisis elaborados que serán expuestos a lo largo de este documento.

El conjunto de datos utilizado en el proyecto ha sido el resultado de haber observado y modificado varios conjuntos de datos que hemos obtenido, lo cual se explicará más adelante.

Agradecimientos

Este proyecto no ha tenido mucha colaboración o apoyo por parte de entidades externas. Sin embargo, debemos agradecer a nuestro tutor, José Alberto Conejero, el seguimiento, guía y ayuda que nos ha brindado en numerosas ocasiones a lo largo del desarrollo del trabajo.

Además, hay que agradecer sin duda a la Universitat Politècnica de València, que nos ha ofrecido software (Statgraphics Centurion, Microsoft Office 365), enseñanza y conexión a Internet, así como a los creadores de las páginas web que nos han ofrecido información y a los creadores de las bases de datos, pues han sido el motor principal de este proyecto y, sin ellos, este trabajo habría estado destinado al fracaso.

Resumen

Este estudio cuyos datos utilizados han provenido fundamentalmente de la PWT (Penn World Table) [1] pretende analizar cual es la situación actual así como la evolución de la cantidad de horas trabajadas. Para resolver este problema se han analizado cuatro aspectos fundamentales: la situación reciente así como su evolución a lo largo del tiempo de las horas trabajadas y otras variables relacionadas (lo que se ha analizado desde un punto de vista global, continental y nacional), su correlación con otras variables y además se ha intentado predecir cual será la situación de las horas trabajadas en un futuro.

En general se observa una tendencia a la baja de las horas trabajadas (aunque la situación es dispar entre países del norte, con tendencia a trabajar menos, y del sur global, con una tendencia mucho más constante), mientras que variables como la productividad, el GDP por hora trabajada y el salario mínimo han presentado un claro aumento de los datos. Por su parte, se ha observado como un mayor número de horas trabajadas no repercute en una mejor economía (medida con el GDP per cápita), ni en una mayor productividad (ocurriendo, de hecho, lo contrario). Finalmente, con los modelos utilizados se ha observado una tendencia a la baja de las horas trabajadas en el mundo.

Palabras clave

Horas trabajadas, productividad, GDP por hora trabajada, salario diario

Índice

1. Introducción	7
1.1. Alcance del proyecto	7
1.1.1. Misión del proyecto	7
1.1.2. Objetivos Generales y Específicos	7
1.1.3. Requisitos y Restricciones del proyecto	8
1.1.4. Asunciones y Supuestos	8
1.1.5. Limitaciones del proyecto	8
1.1.6. Alcance del proyecto	8
1.1.7. Criterios de Éxito y Criterios de Aceptación	8
1.2. Calendario	9
1.3. Presupuesto	10
1.4. Seguimiento y control	11
1.5. Riesgos del proyecto	11
1.5.1. Lista de riesgos	11
1.5.2. Clasificación de los riesgos	11
1.6. Identificación de riesgos	12
1.7. Equipo	12
2. Fuentes, variables y preprocesado de los datos	13
2.1. Fuentes	13
2.2. Modificaciones de los datos	14
2.3. Variables	14
2.4. Selección de casos y análisis previo de los datos	15
2.5. Preparación de los datos	15
2.5.1. Datos faltantes	16
2.5.2. Correlación entre las variables y datos atípicos	17
3. Análisis de los datos	21
3.1. Análisis global	21
3.2. Análisis por continentes	23
3.2.1. ¿Cual era la situación reciente?	23
3.2.2. ¿Cómo han evolucionado las variables?	25
3.3. Análisis por países	27
3.3.1. ¿Cual es la situación recientemente?	27
3.3.2. ¿Cómo han evolucionado las variables?	31
3.4. ¿Son los países más ricos los que más trabajan ?	33
3.5. ¿Aumenta la productividad a medida que disminuyen las horas de trabajo?	34
3.6. ¿Cobran más aquellas personas que más trabajan?	36
3.7. ¿Tienen más calidad de vida aquellas personas que trabajan menos?	36
3.8. Modelo predictivo	37
3.8.1. Predicción global	37
3.8.2. Predicción en España	38
4. Interpretación de los datos	39
5. Estado del arte	40
6. Página Web	40

37.	Gráfico de dispersión: Horas anuales trabajadas/Puntuación en el WHR . .	37
38.	Regresión lineal: horas anuales trabajadas	38
39.	Evolución de las horas trabajadas en España (1950-2019)	38
40.	Regresión lineal: horas trabajadas en España	39

1. Introducción

Este es un mundo donde el ser humano vive ocupado la mayoría del tiempo. Dentro de sus ocupaciones, hay una tarea clara que es vital a lo largo de la vida: el trabajo. Ortega y Gasset ya nos habló de este asunto, y partió de una base muy clara: “El hombre quiere vivir”, y nos explicó que, bajo su punto de vista, hemos desarrollado la técnica en gran medida, lo cual va ligado necesariamente a los conceptos de esfuerzo y profesionalidad. La humanidad ha desarrollado la técnica que nos ha hecho profesionales, adecuados para el trabajo. A medida que el humano ha ido avanzando, se ha hecho evidente la necesidad para la sociedad de trabajar, y ello ha implicado que estemos trabajando una gran parte del tiempo. La humanidad trabaja sin parar. La humanidad vive para trabajar, ya sea por necesidad, por búsqueda de calidad de vida o por pasión por el mismo. Actualmente, se ha generado debate con la cantidad de horas que se trabajan, y han sido muchas las propuestas acerca del asunto, siendo la más destacada la posible reducción de la jornada laboral a cuatro días. Este trabajo nos pondrá en contexto de manera más detallada, y veremos cuánto llegamos a trabajar realmente y qué trayectoria puede llevar esta situación.

1.1. Alcance del proyecto

1.1.1. Misión del proyecto

La misión del proyecto es informar y advertir sobre las horas empleadas en el trabajo y en el futuro y observar las diferencias que hay entre países a nivel mundial (productividad, salud, salario, GDP, IDH).

1.1.2. Objetivos Generales y Específicos

Con respecto a los objetivos generales, este proyecto tiene varias metas:

- Influencia de las horas trabajadas en nuestras vidas.
- Contribución de las vacaciones en la productividad y capacidad de lograr objetivos.
- Relación entre horas de trabajo y GDP de los países.
- Relación entre número de horas trabajadas y salario promedio.
- Relación entre horas de trabajo y el IDH.
- Prever qué tendencia seguirá la cantidad de horas trabajadas.
- Concienciación del tema.

En lo que se refiere a los objetivos específicos (SMART), tiene metas diversas, como, por ejemplo:

- Tratar de simular la posible evolución de un mercado laboral en un contexto pre-pandémico.
- Analizar el posible impacto de la paternidad/maternidad en la productividad del trabajo.
- Abordar en parte el tema de la jornada laboral de 4 días.

1.1.3. Requisitos y Restricciones del proyecto

Este proyecto cuenta con objetivos mínimos, al igual que limitaciones. Algunas restricciones son la existencia de datos faltantes y la necesidad de estudiar su posible tratamiento para mejorar el análisis, así como la garantía de que los datos sean correctos y cumplen el objetivo que buscamos. En cuanto a requisitos, el principal es la creación de nuestra página web. Dicha página web debe facilitar al usuario la orientación en el mercado laboral mediante el análisis de las horas trabajadas y diversas variables como el índice de GPD del país, la productividad o el salario. Esto conlleva necesariamente el éxito en la realización de los análisis y la correcta interpretación e implementación en la página web.

1.1.4. Asunciones y Supuestos

Asumiremos en este proyecto varias proposiciones. En primer lugar, se dará por hecho que un mayor número de horas no necesariamente implican una productividad mayor, pero cabría destacar un posible análisis de si su reducción implica una mayor productividad. Además, no existe evidencia científica contundente acerca de la relación entre las horas trabajadas y las enfermedades.

1.1.5. Limitaciones del proyecto

En este apartado se tratará de las limitaciones del proyecto, es decir, qué haremos y qué no haremos. En lo que respecta al primer apartado, sí crearemos un modelo predictivo cuya función sea el análisis de las horas trabajadas a lo largo de los años, partiendo de datos de fuentes fiables, las cuales corroboramos por su utilidad. Con respecto a lo que no haremos, no nos basaremos en un análisis global sin antes observar los datos faltantes asociados, o no trataremos solamente con promedios de poblaciones estudiadas sin tener en cuenta valores atípicos o valores faltantes en nuestros análisis a realizar.

1.1.6. Alcance del proyecto

El resultado de este proyecto será un estudio en el que se comparará el tiempo de trabajo y la rentabilidad ligada a ello para cotejar la real productividad entre los distintos países, así como un modelo predictivo basado en los análisis para poder vislumbrar de manera orientativa la evolución de las horas que trabajaremos en un futuro. Así pues, pretendemos que el resultado mencionado anteriormente se pueda expresar de forma gráfica mediante análisis estadísticos donde se muestren los ratios evolutivos o involutivos de las poblaciones estudiadas para que el espectador vea de forma mucho más clara la finalidad de nuestros estudios (ya sea mediante la observación de nuestro proyecto o mediante el acceso a nuestra web).

1.1.7. Criterios de Éxito y Criterios de Aceptación

En primer lugar, los criterios de éxito son componentes del proyecto que necesitan colocarse en su lugar para asegurar la realización del de este; así pues, daremos por exitoso nuestro proyecto si y sólo si la web que ofreceremos al público permite a los usuarios orientarse de manera satisfactoria en el mercado laboral y las horas trabajadas. Por otra parte, los criterios de aceptación son el resultado de los requisitos acordados entre el cliente y los creadores del proyecto, respecto a estos, buscaremos la constatación de que los análisis y la información ofrecida son de calidad y fácilmente comprensibles por el público.

1.2. Calendario

Las actividades que en su momento nos planteamos llevar a cabo son las siguientes:

- Análisis y Búsqueda de los datos
 1. Búsqueda datos
 2. Selección y reorganización
 3. Análisis:
 - a) Horas trabajo-productividad
 - b) Horas trabajo-GDP
 - c) Horas trabajo-IDH
 - d) Horas de trabajo-salario
 - e) Horas de trabajo-Word Happiness Report
 - f) Horas trabajadas-Tasa de empleo
 - g) Modelo predictivo
 - h) Contemplar otros posibles análisis
- Estudios y búsqueda de información contextual
 1. Búsqueda y posterior selección de páginas web o estudios
 2. Selección de la información recogida
 3. Diseño y manejo de la información obtenida
 4. Introducir información en los documentos gráficos y web
 5. Contactar con expertos
- Información y Datos
 1. Reestructuración de los datos recogidos
 2. Relación de los datos con los análisis
- Página Web
 1. Creación de la web y de sus elementos básicos (maqueta)
 2. Diseño de la web
 3. Introducción de la información en la web
 4. Revisiones finales
- Documentos Gráficos
 1. Creación del esquema del informe
 2. Memoria (L^AT_EX)
 3. Creación de la presentación (Canva)
 4. Exposición de la presentación

La asignación de las actividades, la secuenciación de estas y la estimación de la duración se pueden observar en la figura 1.

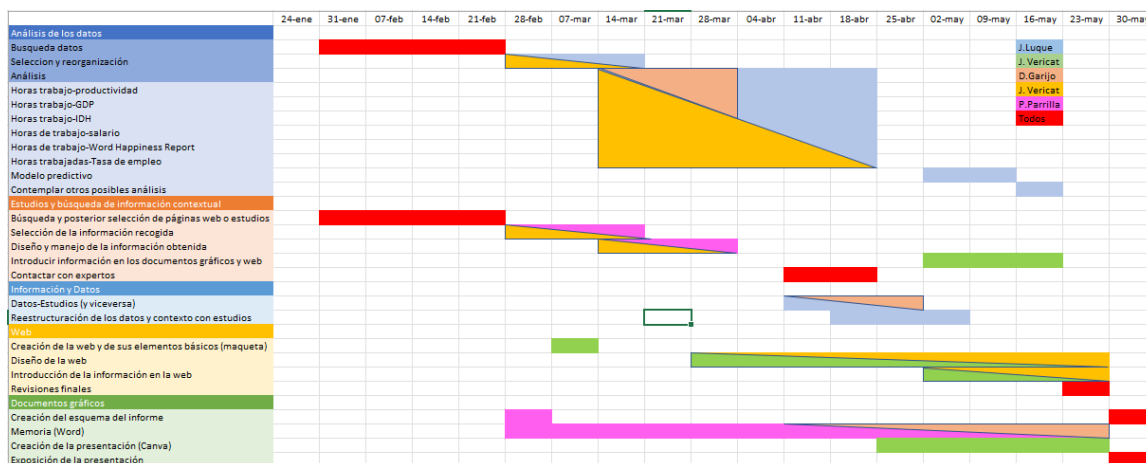


Figura 1: Cronograma del proyecto

Durante el desarrollo del proyecto, nos hemos ceñido bastante al calendario del proyecto, con muy pequeñas variaciones poco relevantes. De esta manera, hemos logrado evitar retrasos, e incluso un avance mayor del esperado, permitiéndonos afrontar las últimas semanas con mayor tranquilidad (pero no por ello menos esfuerzo) y nos ha permitido hacer una revisión más pausada de nuestros análisis y documento gráficos.

1.3. Presupuesto

En este apartado, haremos varias suposiciones acerca del presupuesto necesario para el proyecto. Debido a que este es un proyecto de carácter académico, todos los recursos son de manera tanto indirecta como directamente gratuitos, por lo que se tratará de una estimación de presupuesto real lo más cercana a la realidad posible.

Varios elementos que continuaremos considerando “gratuitos” serán los equipos informáticos (nuestros portátiles) y el paquete Office u otros programas y servicios proporcionados por la Universitat Politècnica de València.

En primer lugar, destacaremos el esfuerzo que invertiremos en el proyecto. El tiempo invertido será de 26 días aproximadamente en durante las clases, lo que hace un tiempo aproximado de 52 horas invertidas durante el horario lectivo. Sin embargo, este proyecto no solo se realizará en clase, sino que es necesario trabajo fuera de la universidad. Asumimos que se invierten aproximadamente 3 horas semanales en este proyecto, con lo que hasta su entrega, hay que sumar a las 52 horas de trabajo en clase, otras 11 semanas de trabajo, lo que implican otras 39 horas más, haciendo un total de 85 horas invertidas en este proyecto de manera real (por persona), a 10€/hora.

En cuanto a los esfuerzos económicos (todos ellos un supuesto), empezaremos por un lugar en el que trabajar, es decir, un estudio. Dicho estudio tendría un precio de unos 400€/mes, siendo 4 meses de trabajo. En segundo lugar, la conexión a Internet de dicho estudio, sobre unos 25€/mes durante 4 meses. En cuanto a licencias, como por ejemplo de un programa para los análisis (en un primer momento se planteo utilizar StatGraphics Centurion pero se acabo usando Plotly, una librería de Python) y un editor web de buena calidad (utilizaremos WebNode, de tipo gratuito), supondremos un gasto total de 40 € en los cuatro meses.

En total, la cifra ascendería a unos supuestos 5990€ y 85 horas de trabajo por persona.

Dicha suma de dinero estaría financiada por el cliente o por la UPV.

1.4. Seguimiento y control

Para hacer un buen seguimiento del proyecto a nivel global, nuestra principal herramienta será el registro de las distintas versiones de todos los archivos de los que iremos disponiendo (hojas de cálculo, presentaciones, memorias, archivos de tipo “.py”, etc.). De esta manera, se podrá controlar la evolución de los mismos. Sus versiones serán guardadas por uno de nosotros.

Este registro de la evolución de los archivos y los datos no solo permite una visión global, sino que, a partir del diagrama de Gantt y las asignaciones de trabajo, se podrá comprobar el progreso individual de las mismas con solo dirigirnos a los archivos y comprobando su actualización y los avances habidos en estos.

Finalmente, y como última medida, más flexible y agradable a priori, es nuestra propia intervención. Nosotros mismos podemos ser nuestra mejor medida de control, pues entre los integrantes del equipo existe una comunicación y un entendimiento fluidos.

1.5. Riesgos del proyecto

1.5.1. Lista de riesgos

Vamos a introducir los riesgos de nuestro proyecto en la siguiente lista:

- Escasez de una tecnología puntera que nos permita analizar los datos sin riesgo al error.
- Problemas existentes en relación con ciertos datos que puedan contener derechos y no nos permitan realizar los análisis planificados.
- La existencia de otros proyectos en los que también estamos involucrados y pueden condicionar la duración y alcance del proyecto.
- La capacidad del sistema para recopilar datos e informes de rendimiento puede ocasionar inconsistencias y restar tiempo de trabajo al equipo de trabajo.

1.5.2. Clasificación de los riesgos

Los riesgos de nuestro proyecto los podemos identificar y dividir en cuatro categorías: técnicos, externos, organizativos y de gestión.

- **Técnicos.** Dentro de la categoría de riesgos técnicos podemos introducir la falta de realización de ciertos análisis debido a una clara escasez de tecnología a nuestra disposición que podría conllevar a un mayor empleo de nuestro tiempo y una reducción del alcance del proyecto (**R1**), así como la realización y modificación de la base de datos existente sin tener bien estudiadas las condiciones tecnológicas en las que nos basamos, donde aun sabiendo lo que tenemos que solucionar para finalizarla (la base de datos), pueden surgir nuevos problemas al utilizar dichos datos para extraer nueva información (**R2**). Relacionado a este riesgo, un ejemplo sería que con un ordenador de mayor capacidad podríamos tener una base de datos mucho más completa que nos proporcionaría un mayor nivel de análisis.

- **Externos.** Por otra parte, al hablar de los riesgos externos de nuestro proyecto podemos hacer referencia a los obvios problemas que encontramos con relación a la existencia de ciertos datos que no podemos emplear en nuestros análisis debido a que estos pueden contener derechos de autor, lo que nos condiciona y obliga a filtrar de más los datos (**R3**). Un ejemplo serían datos provenientes de libros o proyectos que no nos permiten relacionar dichos datos con nuestro trabajo.
- **Organizativos.** Al hablar de los riesgos organizativos hemos de mencionar el hecho de que estamos involucrados en varios proyectos/trabajos en los que buscamos la máxima eficiencia. Así pues, el estar en varios programas al mismo tiempo se nos obliga de ciertas formas a repartirnos el trabajo y nos limita la duración de tiempo que podemos emplearle a este proyecto en particular, así como condicionando el alcance total que podamos tener desde un primer momento si este proyecto lo estuviéramos trabajando de forma independiente (**R4**). En este caso, un ejemplo de riesgo organizativo sería la existencia de un trabajo de Estadística en el que también estaremos involucrados y reducirá nuestras horas de empleo en este proyecto.
- **De gestión.** Por último, los riesgos que denominamos de gestión del proyecto en los que se relaciona nuestra forma de trabajar, y lo que esto conlleva, con la capacidad de los sistemas de los que disponemos para recopilar datos y partes de productividad. Esta relación mencionada anteriormente nos puede ocasionar inconsistencias y disminuir el tiempo de trabajo del equipo en general (**R5**). Un caso de este riesgo, en particular, sería una situación de controversia donde aun teniendo en mente como queremos realizar el análisis de unos datos en específico no sepamos o podamos aplicarlo en el programa informático debido al sistema que empleamos.

1.6. Identificación de riesgos

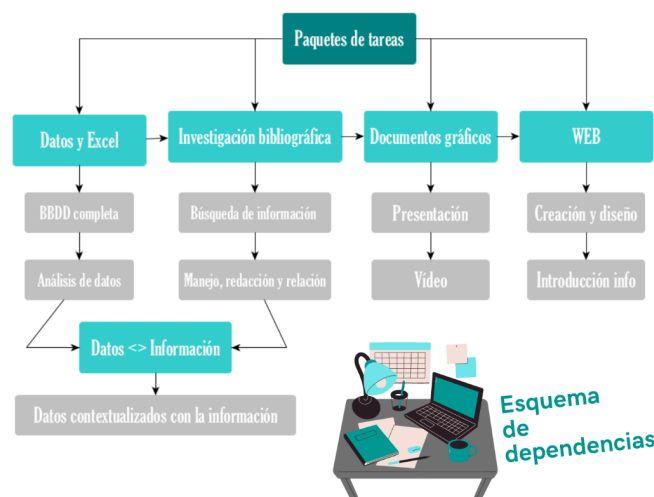


Figura 2: Esquema sobre los riesgos del proyecto

1.7. Equipo

Este equipo está conformado por cinco integrantes: José Valero, Javier Luque, Daniel Garijo, Pablo Parrilla y Júlia Vericat, los cuales, como ya hemos mencionado antes, seremos los encargados de este proyecto acerca de las horas trabajadas, y de responder a las distintas cuestiones que se nos plantean.

La comunicación y la compenetración de la que gozamos serán claves en las relaciones que estableceremos durante el proyecto, logrando un gran ambiente de trabajo que nos permitirá avanzar con destacada rapidez y facilidad. La puesta en común de nuestras variadas habilidades personales, así como las académicas, todas ellas adquiridas en lugares completamente distintos, serán la llave para un estudio de lo más completo, diverso y nacido de la combinación de una gran cantidad de puntos de vista, habilidades, ideas y maneras de proceder. Es así como, nuestro proyecto, en resumen, gozará de amplitud cultural y del consenso, así como los valores ya implícitos en el equipo, como el respeto, la comunicación y la valoración de los distintos puntos de vista y capacidades que producen una gran sinergia entre nosotros y, no solo son vitales en el proyecto, sino que serán potenciadas gracias al mismo.

2. Fuentes, variables y preprocesado de los datos

2.1. Fuentes

Para la realización de este proyecto se han contado con varias bases de datos, siendo la más completa (y por tanto de la cual se han considerado más variables) la que hemos tomado de la Universidad de Groningen, situada en Países Bajos. Esta fue creada para la “Penn World Table” [1], actualizada por última vez el 18 de junio de 2021 a su versión 10.0. Al tener al menos diez versiones y debido al prestigio tanto del evento como de la universidad, podemos afirmar que los datos de los que disponemos son de gran calidad y son una gran referencia en lo que respecta a este proyecto.

De esta base de datos han sido seleccionadas las siguientes variables: `rgdpe` (GDP del lado del gasto), `rgdpe_cap` (GDP per cápita del lado del gasto), `rgdpo` (GDP del lado de la producción), `rgdpo_cap` (GDP per cápita del lado de la producción), `pop` (población en millones), `emp` (población empleada en millones), `hc` (Índice de Capital Humano, que está basado en los años de escolarización), `prod` (productividad, \$/h), `avh` (horas anuales trabajadas) y `avh_week` (horas semanales trabajadas). En este punto hay que comentar varias cuestiones:

- Las variables `rgdpe` y `rgdpo` han sido calculadas teniendo en cuenta las variaciones de precios a lo largo del tiempo y las diferencias de precios entre países, y se miden en dólares internacionales en precios de 2011.
- Las variables `rgdpe_cap` y `rgdpo_cap` han sido calculadas únicamente dividiendo `rgdpe` y `rgdpo` entre la población del respectivo año.
- La variable `avh_week` ha sido calculada dividiendo las horas anuales trabajadas (`avh`) entre 52 semanas. [2]

Además de esta fuente para realizar un análisis más completo ha sido necesario considerar variables que no se tenían en cuenta en la PWT para lo cual se ha tenido que buscar más bases de datos. Estas son las siguientes:

- **OECD** [3], de donde se ha seleccionado únicamente la variable `GDP_hour` (GDP por hora trabajada).
- **Our World in Data** [4] (aunque la fuente original es PovCalnet cuyos datos provienen del Banco Mundial) de donde se ha seleccionado la variable `inc_con` (salario mediano diario), datos que han sido obtenidos mediante encuestas (dependiendo del país preguntando por los ingresos del hogar o por los gastos del mismo). Estos datos

se ajustan para tener en cuenta las variaciones de los precios a lo largo del tiempo y las diferencias de precios entre países, y además, se expresan en dólares internacionales.

- **ScienceDirect** [5], de donde se han obtenido los datos de days_vac (días de vacaciones y festivos nacionales al año).
- **United Nations Development Programme** [6] , de donde se han obtenido los datos de idh (Índice de Desarrollo Humano).
- **Kaggle** [7], de donde se han obtenido los datos de happ (puntuación en el World Happiness Report), así como variables relacionadas, tales como eco (GDP per capita), life_exp (esperanza de vida), freed (libertad), trust (confianza en el gobierno) y gen (generosidad).

2.2. Modificaciones de los datos

A partir de todas las fuentes comentadas se creó una gran base de datos con la que poder realizar los análisis. En esta, además de todas las mencionadas variables, se crearon dos variables más: UT, que sirve como filtro para poder realizar los análisis (esto se comentará en profundidad en la sección de Selección de Casos) y continent (que sirve para poder analizar la situación a nivel continental).

2.3. Variables

Variable	Descripción	Tipo	Rango	Tmp med
country	Nom.País	Cualitativa	-	-
continent	Nom.Cont.	Cualitativa	-	-
year	Año	-	1950-2019	1950-2019
UT	0=no considerado 1=considerado	Cualitativa	-	-
avh	Promedio de horas anuales trabajadas/personas ocupadas	Cualitativa	-	-
avh_week	Promedio de horas anuales trabajadas por semana	Cuantitativa continua	26,55-58,45	1950-2019
rgdpe	GDP real del lado del gasto (en millones de US\$ de 2011)	Cuantitativa Continua	20,36-20860506	1950-2019
rgdpe_cap	GPD real del lado del gasto per capita	Cuantitativa Continua	244,601-283541,2	1950-2019
rgdpo	GDP real del lado de la producción (en millones de US\$ de 2011)	Cuantitativa continua	27,65-20595844	1950-2019
rgdpo_cap	GDP real del lado de la producción per capita	Cuantitativa continua	214,55-444763,3	1950-2019
pop	Población(millones)	Cuantitativa Continua	0,0044-1433,78	1950-2019
emp	Número de personas contratadas (en millones)	Cuantitativa continua	0,0012-799,30	1950-2019
hc	Índice de capital humano, basado en años de escolaridad y retornos a la educación	Cuantitativa continua	1-4,35	1950-2019

inc_con	Percentil 50 (mediana): nivel de ingresos o consumo diario	Cuantitativa continua	0,45-206,27	1870-2020
days_vac	Días de vacaciones y festivos para los trabajadores de la producción a tiempo completo en actividades no agrícolas	Cuantitativa discreta	4- 42,5	1870-2000
prod	Productividad	Cuantitativa continua	0,58-109,77	1870-2020
idh	Índice de desarrollo humano	Cuantitativa Continua	0,192-0,957	1990-2019
happ	Posición en el World Happiness Report	Cuantitativa continua	2,693-7,769	2015-2019
GDP	PIB por hora trabajada	Cuantitativa continua	2,99-104,64	1970-2019
eco	GDP per capita	Cuantitativa continua	0,0153-2,096	2015-2019
life_exp	Esperanza de vida	Cuantitativa continua	0,006-1,141	2015-2019
freed	Libertad	Cuantitativa continua	0,00589-0,724	2015-2019
trust	Confianza en el Gobierno	Cuantitativa continua	0,001-0,55191	2015-2019
gen	Generosidad	Cuantitativa continua	0,002-0,838	2015-2019

2.4. Selección de casos y análisis previo de los datos

A lo largo de este estudio, se puede observar como existen países que no han sido considerados para hacer algunos gráficos (aquellos en las que un exceso de casos impedía la correcta visualización del mismo). Esto es debido, además de por la cuestión inmediatamente comentada, fundamentalmente a dos factores: la existencia de muchos datos faltantes (como se verá en el siguiente punto), y a que pueden ser países poco atractivos para un público occidental. Sin embargo, para el análisis global y por continentes, han sido considerados todos los países con datos en la variable estudiada.

Cabe destacar, además, que para la selección de casos ha primado siempre el criterio estadístico, por encima del criterio subjetivo. Prueba de esto es que sí han sido considerados países como Camboya, Myanmar, Bangladesh, Singapur... puesto que previo a dejar de considerar algunos países se hizo un primer análisis descriptivo con todos los casos para asegurar que la calidad del estudio no se viera influenciada por factores culturales.

Técnicamente, para filtrar los países se ha utilizado la variable UT anteriormente comentada.

2.5. Preparación de los datos

Además, del comentado análisis gráfico de las variables, en este estudio, antes de realizar ningún tipo de análisis concreto, se ha realizado una análisis general de los datos, análisis que incluye la obtención de varios parámetros muestrales de las diversas variables, estudio de los datos faltantes y de su posible imputación, análisis de datos atípicos, estudio de las correlaciones...

2.5.1. Datos faltantes

En primer lugar, con el objetivo de poder visualizar gráficamente cómo es la base de datos usada, se ha utilizado un mapa de calor (observable en la Figura 3) en el cual los datos faltantes quedan marcados como celdas blancas y, por el contrario, las celdas negras representan datos (no faltantes).

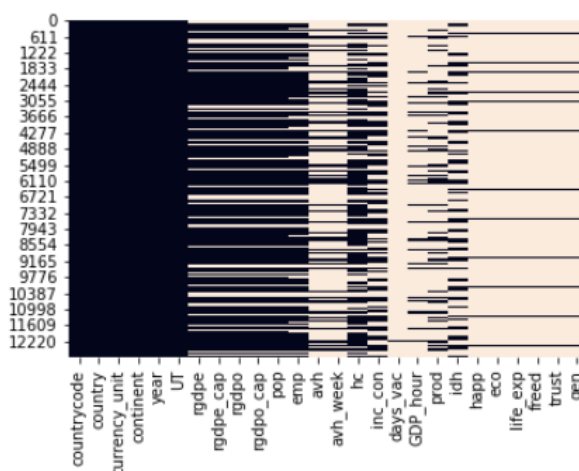


Figura 3: Análisis de los datos faltantes

En el gráfico, se observa como la cantidad de datos faltantes es muy distinta en cada variable: por ejemplo: rgdpe, rgdpe_cap, rgdpo, rgdpo_cap, pop y emp no presentan casi datos faltantes. Por el contrario variables como avh , avh_week, hc e inc_con presentan un nivel de datos faltantes medio. Finalmente, entre otras, variables como days_vac (dias de vacaciones al año), happ (puntuación en el World Hapiness Report), life_exp (esperanza de vida) presentan muchos datos faltantes.

En relación a esto, en este estudio finalmente se decidió no tratar los datos faltantes debido fundamentalmente a que estos se concentran en países que, por lo general, no han sido considerados. Sin embargo, pese a esta gran cantidad de datos faltantes, como se podrá observar posteriormente, los gráficos que se han obtenido son muy completos. Sin embargo, es cierto que este factor tiene un impacto mayor en el análisis por continentes, pues sobre todo en el caso de África y Oceanía, sus países cuentan con muy pocos datos en algunas variables, lo que al obtener la media puede distorsionar los resultados (esto que verá más en detalle en próximos apartados).

Pese a esto, antes de dejar de considerar estos países, se han probado diferentes técnicas de imputación para datos faltantes:

- Establecer a todos los casos con datos faltantes la media de la respectiva variable aleatoria.
- Se utilizó el algoritmo KNN (k-nearest neighbors algorithm) obteniendo unos resultados muy similares a los del caso anterior debido a que, al haber tantos datos faltantes, el valor obtenido acababa siendo la media de la variable.
- Técnicas de interpolación, obteniendo resultados algo más prometedores que en los casos anteriores, pero con algunas cuestiones de dudosa fiabilidad estadística.

2.5.2. Correlación entre las variables y datos atípicos

Para poder hacer los gráficos concretos, concretamente los de dispersión por países y posibles modelos predictivos, es necesario conocer el grado de relación entre las variables. Para ello, se ha obtenido la matriz de correlación (Figura 4) resultante de calcular el coeficiente de correlación de Pearson

$$r_{xy} = \frac{Cov_{xy}}{S_x S_y} = \frac{S_{xy}}{S_x S_y}$$

para cada par de variables, de forma que si este parámetro es próximo a 1 o a -1 existe relación entre estas (directa e inversa respectivamente), y si, por el contrario, este está próximo a 0, las variables no presentan mucha relación lineal entre ellas.

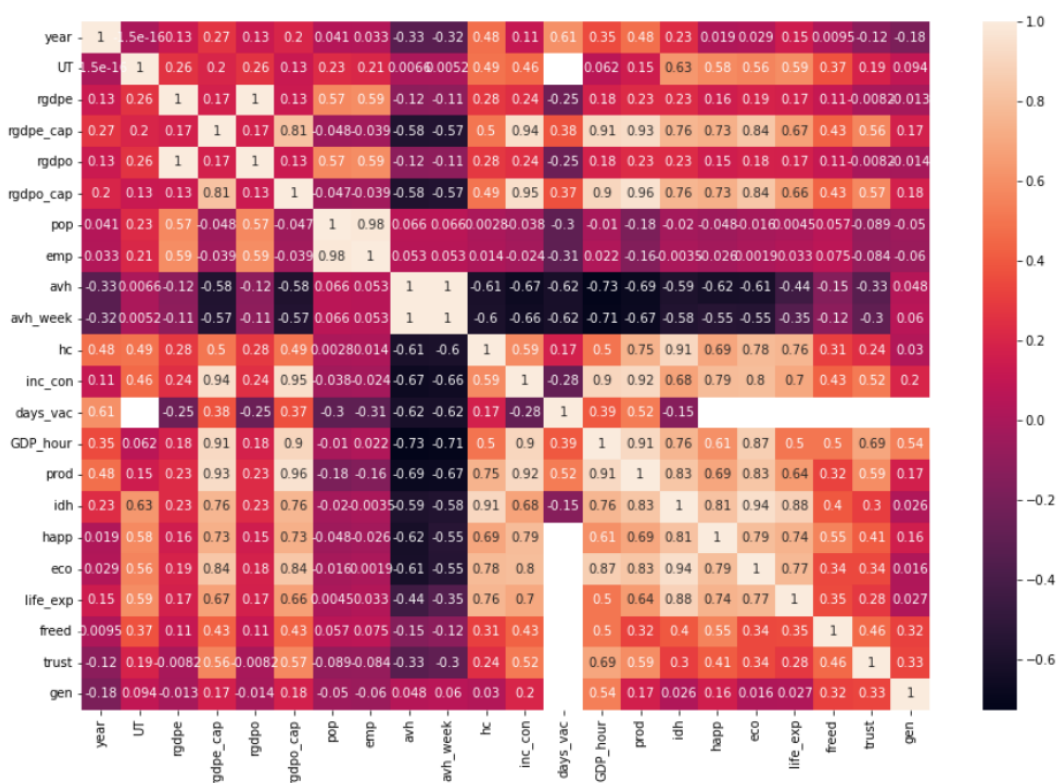


Figura 4: Matriz de correlación

Mirando la figura 4, concretamente la fila de horas trabajadas (avh), que es el eje principal del proyecto (y análogamente mirando avh_week), se observa que esta variable tiene una fuerte relación con variables como: el Human Capital Index (hc), el salario medio diario (inc_con), los días de vacaciones al año (days_vac), el GDP por hora trabajada (GDP_hour), la productividad (prod), el Índice de Desarrollo Humano (idh) y la valoración en el World Happiness Report (happ). Por otro lado, hay otras variables con algo menos de relación como la esperanza de vida (life_exp) o la confianza en el gobierno (trust).

Sin embargo, hay otras variables cuya relación con avh es muy baja, entre las cuales hemos que resaltar cuatro que tienen un grado de relación extremadamente bajo con avh. Estas variables son rgdpo y rgdpe (que a priori pueden parecer que están relacionadas con las horas de trabajo, y así es) y, pop y emp. En este punto es muy importante hablar de los

datos atípicos, que son especialmente extremos en estos casos. Para analizar esta cuestión en las figuras 5, 6, 7, 8 se muestran los gráficos de dispersión entre avh y las variables mencionadas anteriormente en el año 2019.



Figura 5: Gráfico de Dispersión: Horas anuales trabajadas/GDP per lado de la producción

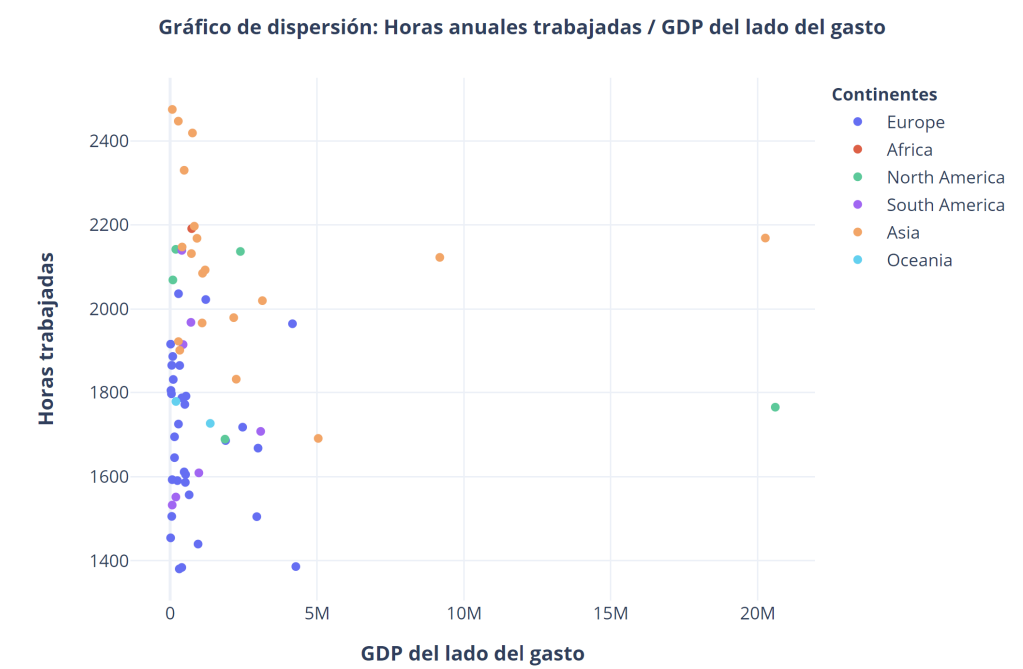


Figura 6: Gráfico de Dispersión: Horas anuales trabajadas/GDP per lado del gasto

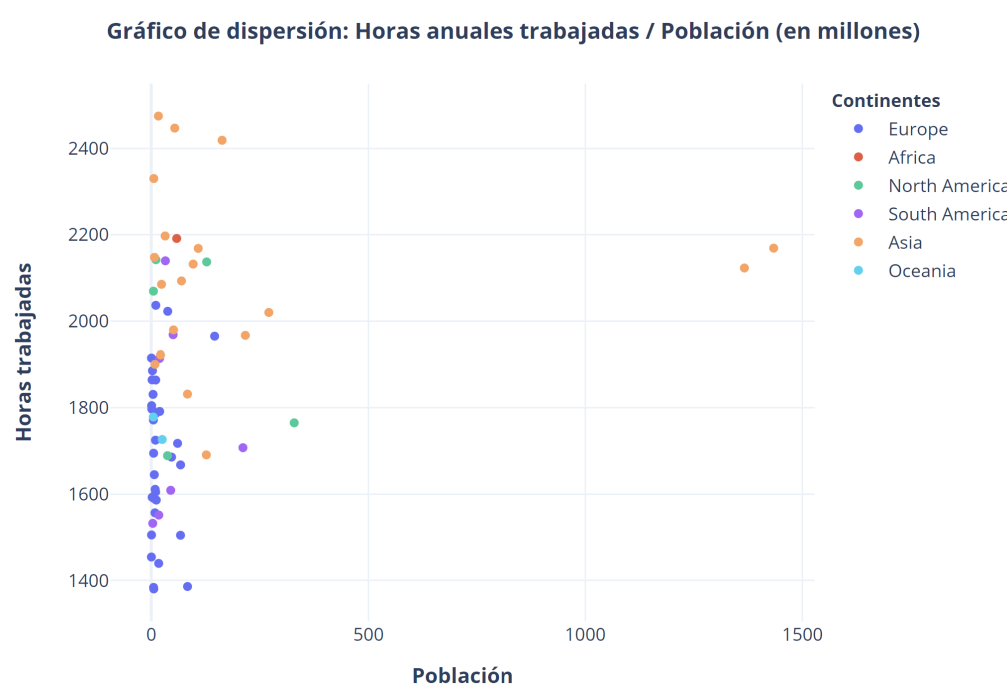


Figura 7: Gráfico de Dispersión: Horas anuales trabajadas/Población (en millones)

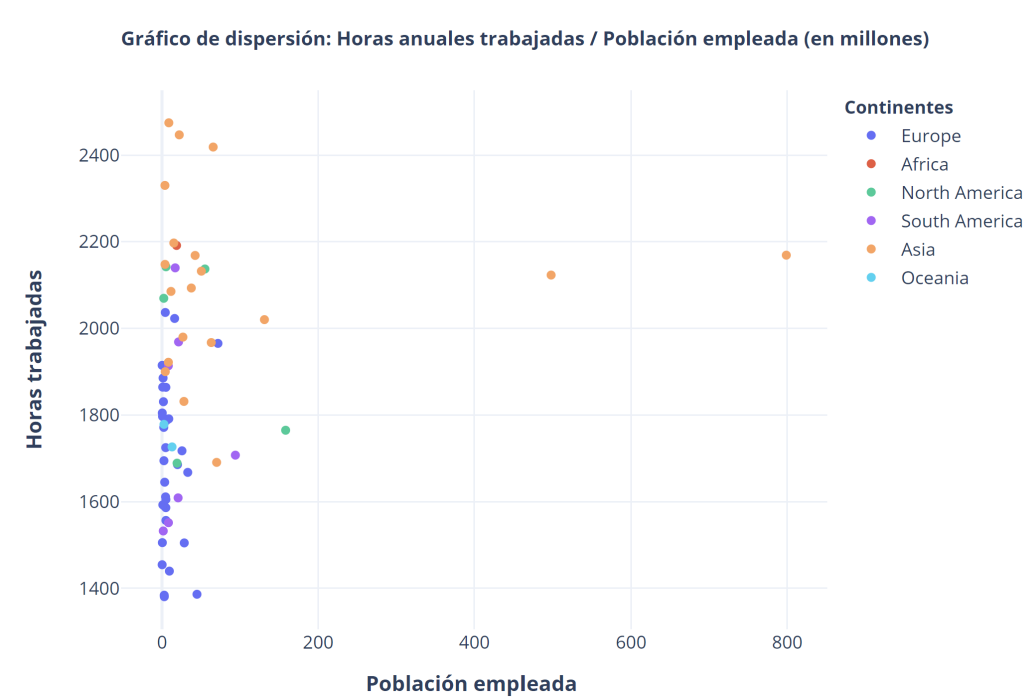


Figura 8: Gráfico de Dispersión: Horas anuales trabajadas/Población empleada (en millones)

En estos gráficos se pueden observar claramente varios datos atípicos:

- En el primer y segundo gráfico, se observa como China y Estados Unidos fundamentalmente presentan un GDP del lado de la producción y del gasto mucho mayor al resto. Se observa además, como India también destaca (aunque en mucha menor medida) del resto de países.
- Por su parte, en el tercer y cuarto gráfico, salta a la vista como India y China presentan mucha más población (tanto en general, como personas contratadas) que el resto de los países.

Por tanto, para poder observar si estas variables están relacionadas con las horas de trabajo anuales, hay que dejar de considerar estos países (Estados Unidos, China e India). La matriz de correlación en este segundo caso se puede observar en la figura 9.

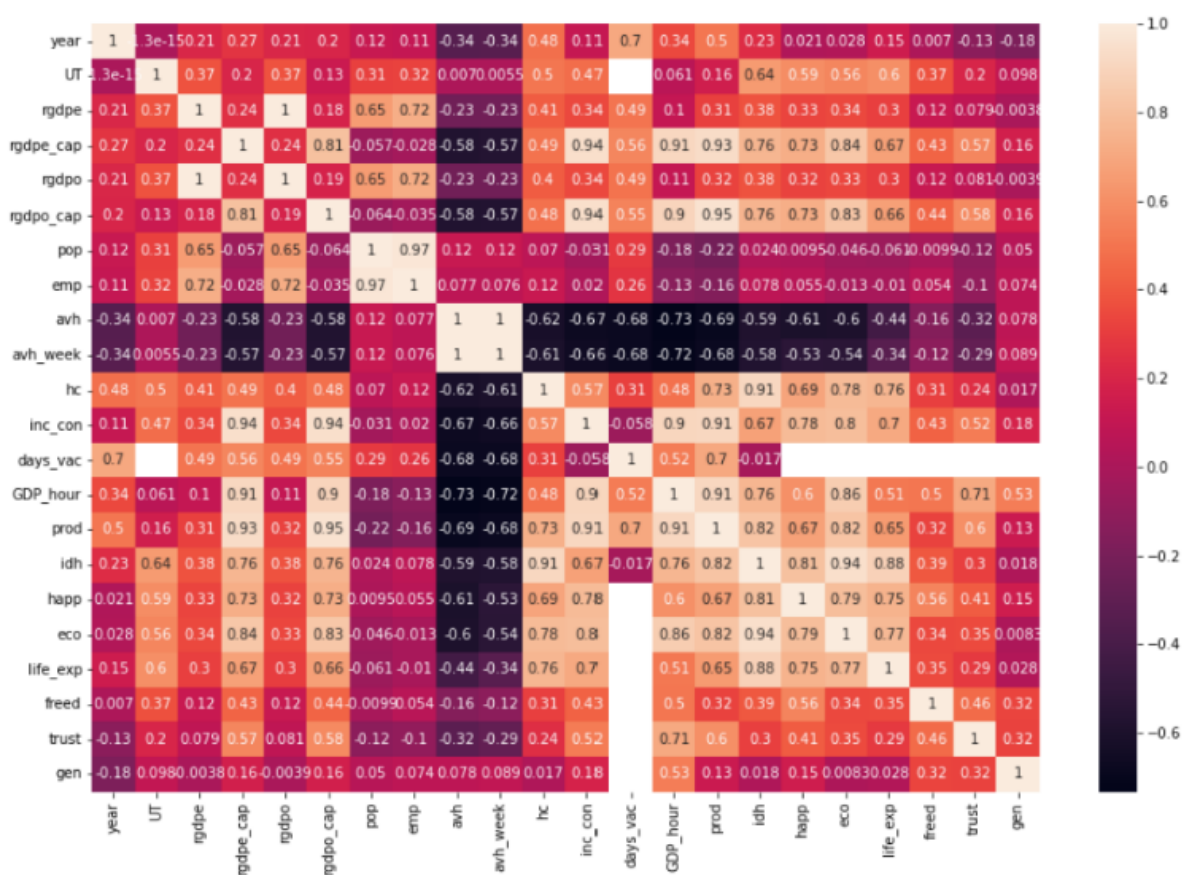


Figura 9: Matriz de correlación (sin considerar EEUU, China e India)

En la matriz se observa que, pese a que el coeficiente de correlación ha aumentado en todos los casos, siguen siendo valores muy lejanos a 1 (lo que indica que estadísticamente estas variables no tienen una fuerte relación con las horas trabajadas). Pese a esto, hay que ser cautos con respecto de las variables rgdpo y rgdpe puesto que, como se puede observar, la situación es bien distinta cuando se habla de valores per cápita, mostrando que verdaderamente estas variables sí que están relacionadas con avh. Es por esto por lo que en este estudio nunca se utilizarán los valores de rgdpo y rgdpe como tal, sino que se utilizarán sus valores per cápita.

Sin embargo, la situación es distinta en el caso de freed (libertad), trust (confianza en el gobierno) y gen (generosidad), que tienen una varianza muy baja, por lo que verdaderamente no tienen relación con avh. Por tanto, estas variables pese a haber sido elegidas para diseñar la base de datos por ser variables del World Happiness Report, no serán analizadas en profundidad en el estudio ya que, precisamente por lo comentado, ya se tienen en parte en cuenta en la variable happ.

3. Análisis de los datos

3.1. Análisis global

En primer lugar, para poder hacer análisis más concretos, se plantea un análisis global que permita comprender cuál es la situación general en el mundo. Para ello, se ha calculado la media de todas las variables, obteniendo los siguientes resultados.

Por lo que respecta a las horas anuales trabajadas, en la figura 10, se puede observar cómo en el mundo la media de estas ha tenido una tendencia claramente bajista desde 1950, pues los datos disminuyeron casi un 15 % de 2171 horas anuales (41.75 horas semanales) en 1950 a 1849 horas anuales (35.55 horas semanales) en 2019.

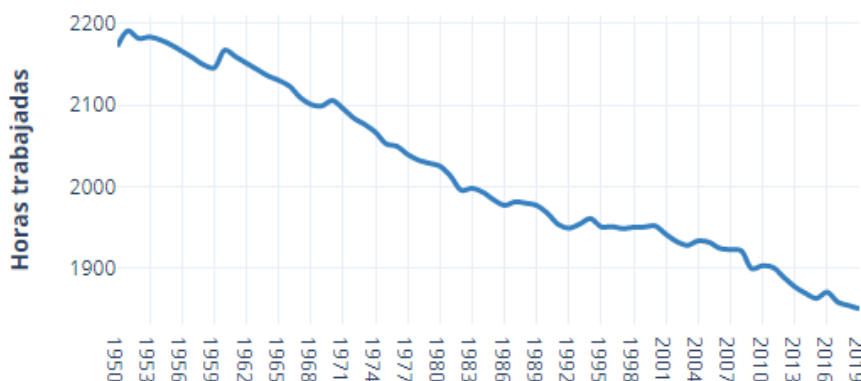


Figura 10: Evolución de las horas anuales trabajadas (1950-2019)

Por contra, la evolución ha sido bien distinta en el caso de la productividad (\$/h), el GDP por hora trabajada y el salario medio diario:

- La productividad ha aumentado de forma prácticamente constante desde 1950 hasta 2019 salvo en dos momentos: el año 1969 (debido fundamentalmente a un descenso en los datos de Oceanía y Asia) y en 2008 (año en el que comenzó una gran crisis bursátil que estalló con la caída del banco estadounidense Lehman Brothers pero que ya estaba fraguándose desde el año 2007 [8]). Pese a esto, los datos aumentaron casi un 384 % de 7.164 \$/h en 1950 a 34.64 \$/h en 2019. Esta evolución se puede observar en la figura 11.

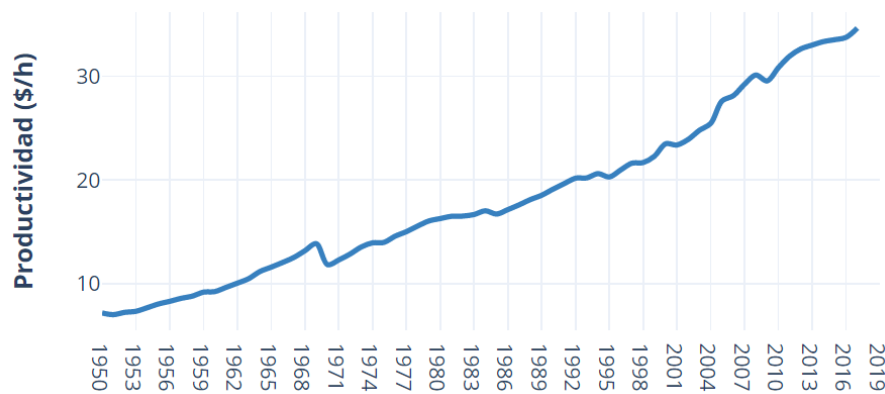


Figura 11: Evolución de la productividad (1950-2017)

- Como se puede observar en la figura 12, el GDP por hora trabajada también ha aumentado progresivamente desde 1970 hasta 2017, presentando, no obstante, algunos periodos de estancamiento o incluso de disminución de los datos (como por ejemplo en 1990 y en 2008). Con esto, los datos aumentaron aproximadamente un 107 % entre estos dos años.

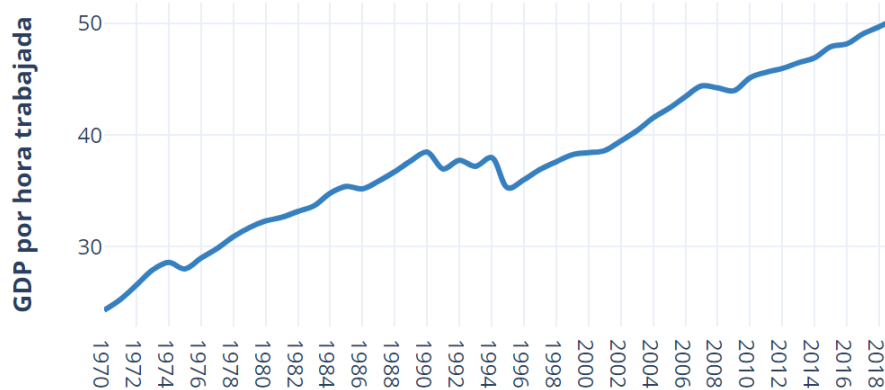


Figura 12: Evolución del GDP por hora trabajada en el mundo (1970-2019)

- Por su parte, como se puede ver en la figura 13, el salario diario también ha mostrado un comportamiento generalmente al alza desde 1981 hasta 2019, presentando algunos ligeros descensos en los mismos años que en el caso anterior. En este caso, la tasa de variación entre 1981 (con un salario promedio de 10\$/día) y 2019 (con aproximadamente 16 \$/día) es del 60 %.

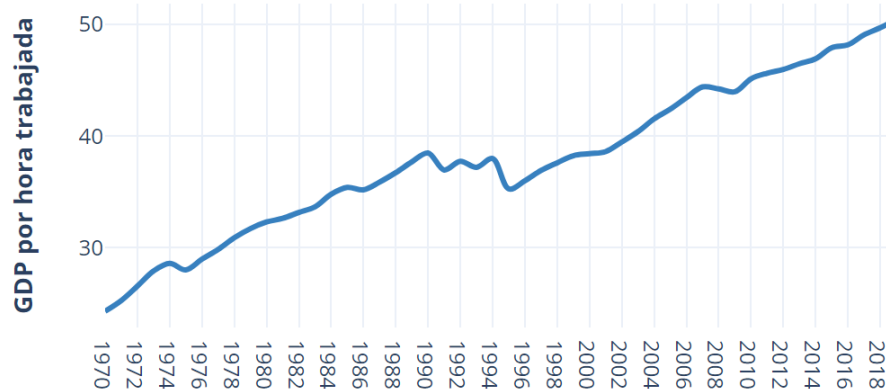


Figura 13: Evolución del salario mediano diario en el mundo (1970-2019)

3.2. Análisis por continentes

Centrando más el foco, también se puede analizar la situación a nivel continental. En este punto, hay que ser cautos con los resultados puesto que los datos han sido obtenidos calculando la media de cada continente para cada variable, por tanto, en el caso de África y Oceanía (donde únicamente dos países tienen datos: Nigeria y Sud África, y Australia y Nueva Zelanda) los resultados pueden ser más altos que en otros continentes en los cuales algunos países pueden compensar a otros. Teniendo esto presente, se pueden realizar diversos análisis.

3.2.1. ¿Cual era la situación reciente?

Por lo que respecta a las horas anuales trabajadas en la figura 14 se observa como son África (2191 h/año o 42.13 h/semana) y Asia (2110 h/año o 40.57 h/semana) los dos continentes en los cuales las personas trabajaron más horas en 2019, y seguidamente se encuentra Norte América (1960 h/año o 37.69 h/semana).

Por detrás, Sud América (1774 h/año o 34.11 h/semana) y Oceanía (1752 h/año o 33.69 h/semana) presentan datos muy similares, y finalmente, Europa es el continente en el cual se trabajó menos horas en dicho año, con una media de 1689 horas por trabajador, es decir, 32.48 horas a la semana.

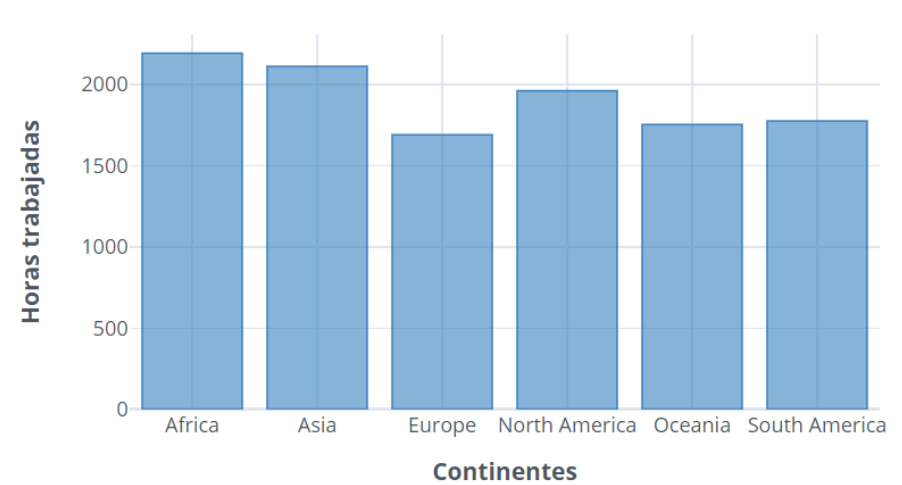


Figura 14: Horas anuales trabajadas por continentes (2019)

En cuanto al resto de variables aleatorias intimamente relacionadas con las horas trabajadas, la situación estos últimos años es la siguiente:

- **Productividad.** En la figura 15 se observa como los dos continentes que presentaron una productividad (\$/h) mayor en 2017 son Oceanía (46.69 \$/h) y Europa (46.59 \$/h). Algo por detrás, se encuentra Norte America con una productividad de (37.33 \$/h), y bastante por detrás, se encuentran Asia (21.2 \$/h), Sud América(18 \$/h) y África (11.77 \$/h).

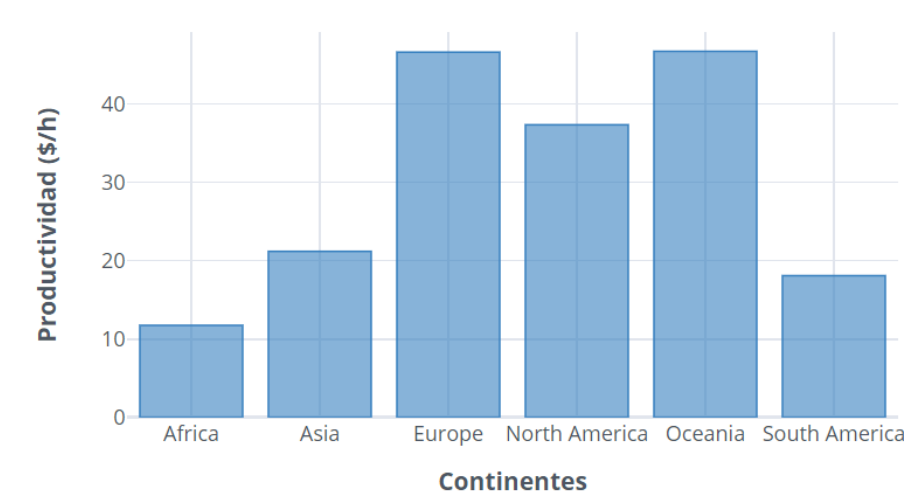


Figura 15: Productividad (\$/h) por continentes (2017)

- **GDP por hora trabajada.** En la figura 16 se observa como, en la línea de lo anterior, el continente con un mayor GDP por hora trabajada en 2019 fue Europa (55.48\$), seguido muy de cerca por Oceanía (48.49\$), Asia (44.85\$) y Norte América (41.58\$). Bastante por detrás, se encuentran Sudamérica (18.1\$) y África (11.77\$).

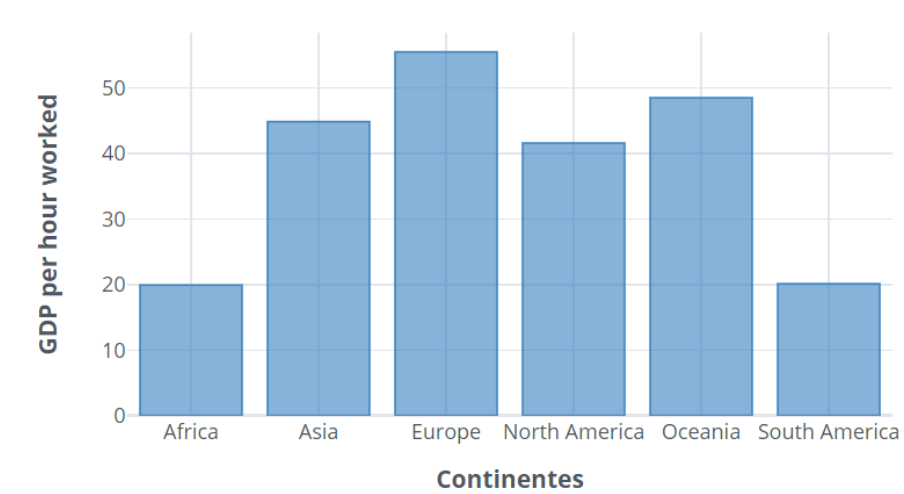


Figura 16: GDP por hora trabajada por continentes (2019)

- **Salario diario.** En la figura 17 se observa como, en este caso, la situación es bastante similar al caso anterior, pues Europa es el continente donde los trabajadores obtuvieron un mayor salario diario en 2019 con un total de 32.14 \$/día. Por detrás, con datos muy similares, se encuentran Oceanía, Asia y Norte América; y bastantes por detrás del resto se encuentran Sud América y África.

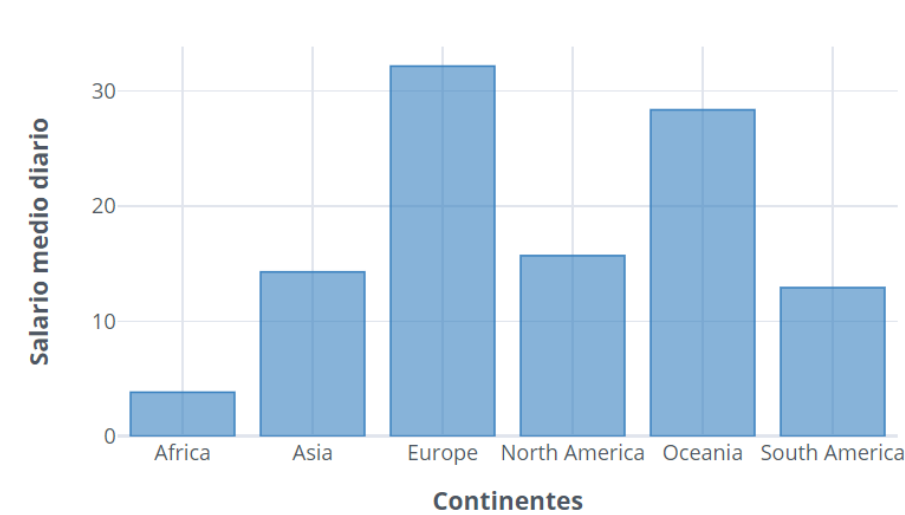


Figura 17: Salario mediano diario por continentes (2019)

3.2.2. ¿Cómo han evolucionado las variables?

En primer lugar, por lo que respecta a las horas anuales trabajadas, en la figura 18 se puede observar una tendencia generalmente descendente en todos los continentes. Particularmente la situación en cada uno es la siguiente:

- **África.** Se puede ver un comportamiento a la baja desde 2005 hasta 2017 (obviando el dato faltante de 2016), y un incremento en 2018. Con esto, en el continente africano, las personas trabajaron casi un 12 % menos en 2019 que en el año 2001.
- **Asia.** En primer lugar, se observa un incremento de las horas trabajadas desde 1950 hasta 1960, posteriormente, se produjo una cierta estabilización de los datos hasta 1969. En este punto, parece que los valores se reducen mucho, pero hay que tener cuidado, pues es en 1970 el año en que el número de países con datos aumenta, lo que hace reducir la media del continente. A partir de ese momento, los datos han ido descendiendo paulatinamente. Con esto, en 2019, los asiáticos trabajaron casi un 8.38 % menos que en 1970.
- **Europa.** En el continente europeo se observa un descenso prácticamente constante desde 1950, salvando el periodo de 1993 a 1995 cuando aumentaron ligeramente los datos. Así, los europeos trabajaron en 2019 un 22 % menos que en 1950.
- **Norte América.** En este caso, también se puede ver, salvando el año 1987 (año en el que comienzan a aparecer los datos de Costa Rica, siendo este el país norteamericano con más horas anuales trabajadas) una tendencia irregular, aunque con inclinación a la disminución de las horas trabajadas
- **Oceanía.** Se puede observar un comportamiento a la baja prácticamente en todo momento, salvando algún aumento de los datos puntual. Con esto, en Oceanía, las personas trabajaron en 2019 un 19 % menos de lo que lo hicieron en 1950.
- **Sud América.** En este caso, se puede ver una tendencia muy similar a la del caso anterior, ya que los sudamericanos trabajaron casi un 20 % menos en 2019 que en 1950.

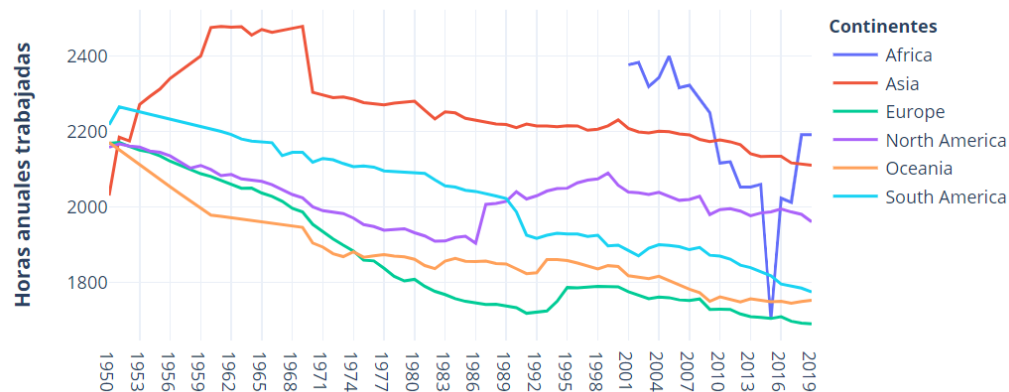


Figura 18: Evolución de las horas anuales trabajadas por continentes (1950-2019)

Por lo que respecta al resto de variables intimamente relacionadas con las horas trabajadas, su evolución fue la siguiente:

- **Productividad.** En la figura 19, se puede observar como, en general, todos los continentes presentan una tendencia ascendente con algunas épocas de estancamiento y decrecimiento de los datos (aunque en general estos son debidos a que en los años donde los datos decrecen empiezan a haber datos de países que antes no tenían).

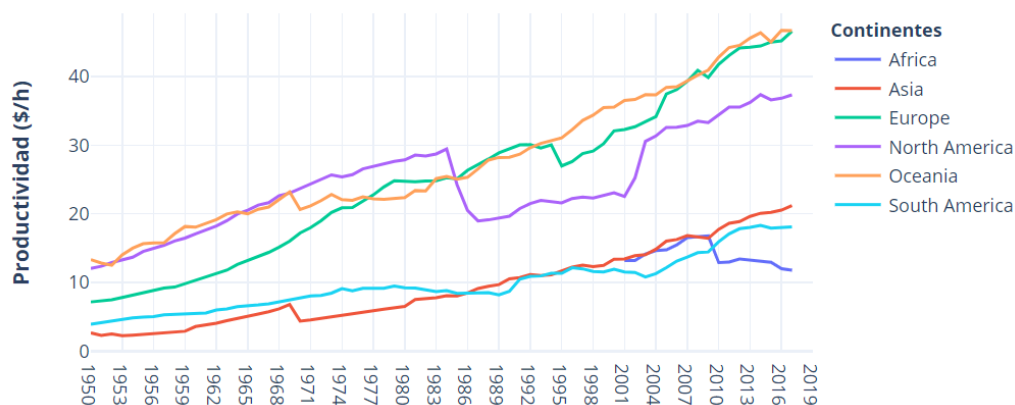


Figura 19: Evolución de la productividad por continentes (1950-2017)

Las variaciones de algunos continentes son las siguientes: los europeos fueron en 2017 casi un 552 % más productivos que en 1950, los trabajadores asiáticos en 2017 fueron un casi un 703.03 % más productivos que en 1950, los norteamericanos fueron en 2017 casi un 212 % más productivos que en 1950...

- **GDP por hora trabajada.** En la figura 20, se puede observar como, de la misma forma que en el caso anterior, todos los continentes presentan una tendencia al alza (salvando la situación comentada anteriormente).

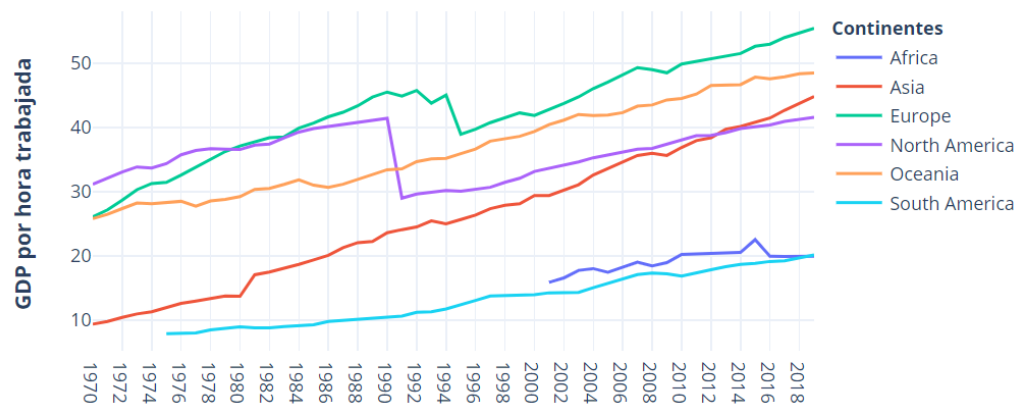


Figura 20: Evolución del GDP por hora trabajada por continentes (1970-2019)

- **Salario mediano diario.** En la línea de lo anterior, en la figura 20, se puede observar como todos los continentes, salvo África cuyos datos se han mantenido prácticamente constantes, presentan una tendencia creciente.

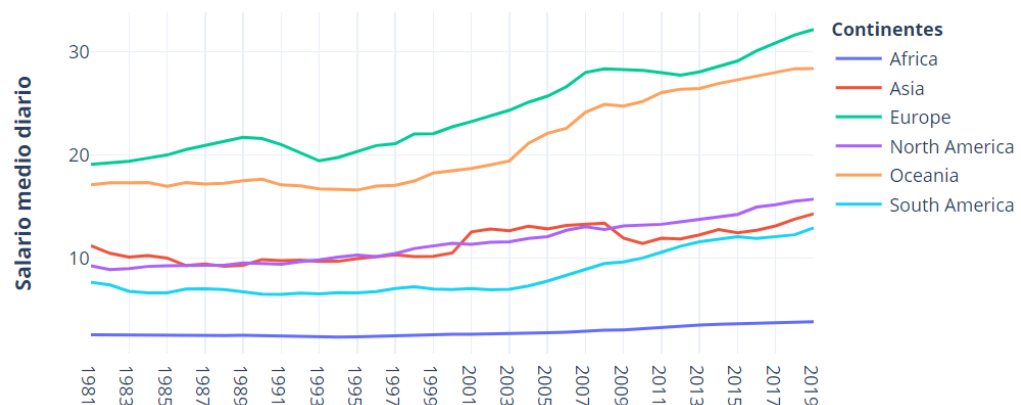


Figura 21: Evolución del salario mediano diario por continentes (1981-2019)

3.3. Análisis por países

3.3.1. ¿Cual es la situación recientemente?

Por lo que respecta a las horas trabajadas, como se puede observar en la figura 22, la situación es tremendamente dispar en distintos países: aquellos que más horas trabajaron en 2019 fueron fundamentalmente países asiáticos como Camboya (2474 horas al año, es decir, unas 47.57 horas semanales), Birmania (Myanmar), Bangladesh, Singapur, Malasia...; mientras que los países donde los trabajadores trabajaron menos horas en ese mismo año fueron principalmente países europeos como Países Bajos, Alemania, Noruega y Dinamarca, siendo este último el país donde menos horas se trabajaron en 2019 (1380 horas al año, unas 26 horas semanales).

En cuanto al continente europeo, los países donde los empleados trabajaron más horas al año fueron Grecia (2036 horas anuales, 39 horas semanales), Polonia, Rusia, Malta... Otras cuestiones a destacar son, por ejemplo, que Japón es el país de Asia donde menos horas se trabajó en 2019 con unas 1691 horas anuales, que en Canadá (1689 h/año) se trabajaron

menos horas que en Estados Unidos (1765 h/año), y que España es el decimonoveno país del mundo donde menos se trabajó en 2019.

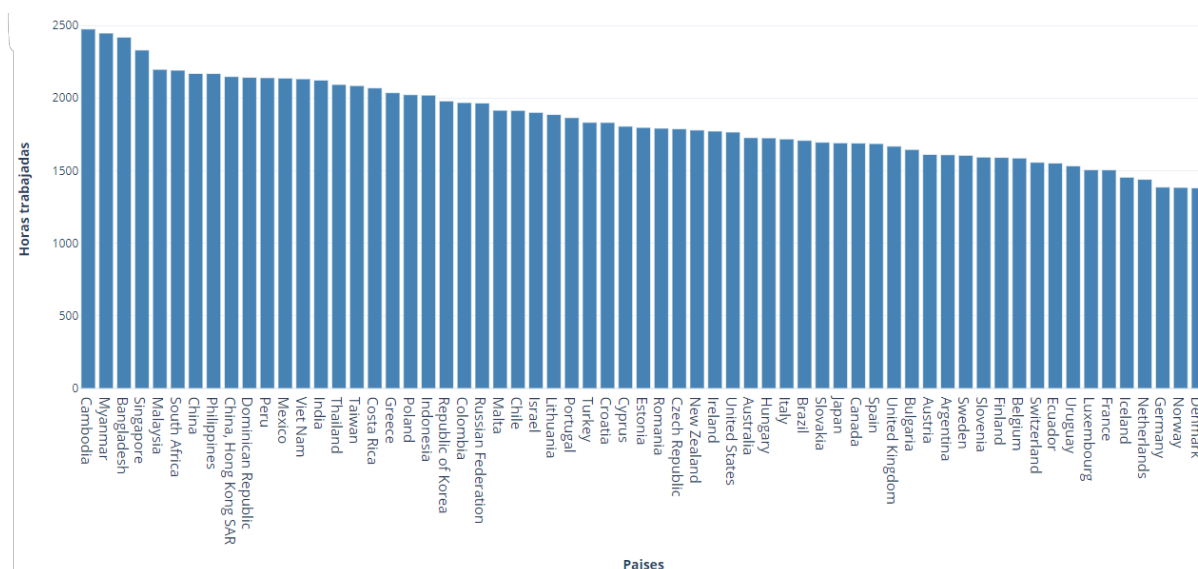


Figura 22: Horas anuales trabajadas por países (2019)

Por su parte la situación del resto de variables es la siguiente:

- **Productividad.** En este caso, como se puede observar en la figura 22 la situación es bien distinta, pues precisamente los países donde más horas se trabaja son aquellos que presentaron una productividad menor en 2017, como por ejemplo, que Camboya presenta un productividad de 2.24 \$/hora. En el extremo opuesto, Irlanda (con una productividad de 99.13 \$/h), seguido de Noruega (80.83\$/h) son los países con una mayor productividad. Otras cuestiones a comentar son:
 - Estados Unidos es más productivo que Canadá, siendo estos dos países (con diferencia los dos países con mejores cifras de América del Norte).
 - España se encuentra a media tabla respecto a los países de Europa con una productividad de 51.17 \$/h, siendo superada por países de su entorno como Italia (53.29 \$/h), Francia (62.79 \$/h) o Alemania (66.71 \$/h); sin embargo está muy por encima de países como Grecia (30.18 \$/h) o Portugal (27.8 \$/h).

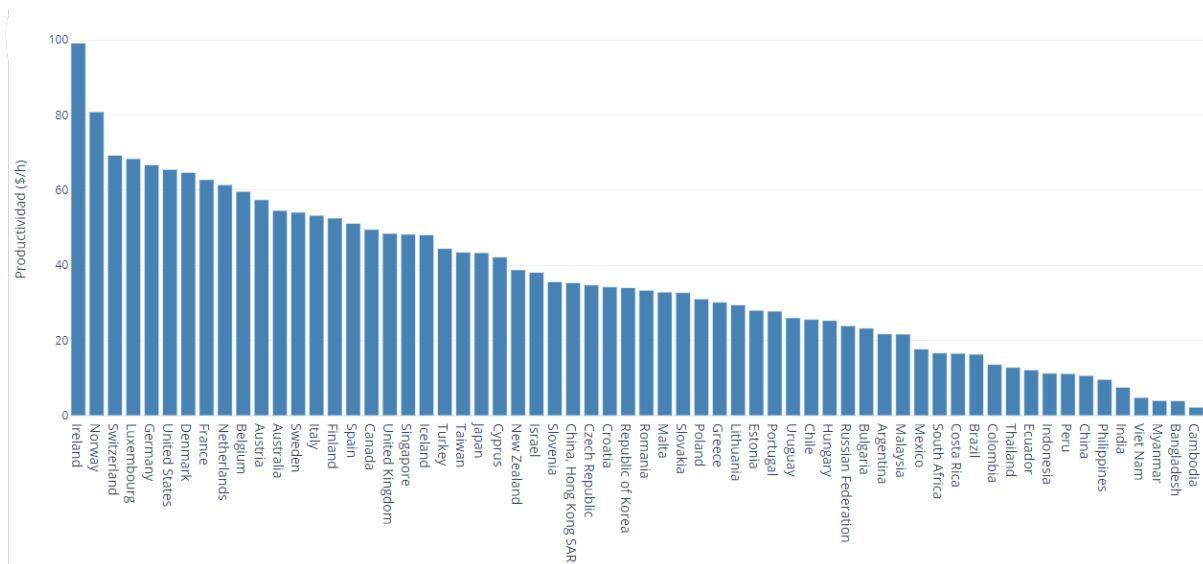


Figura 23: Productividad (\$/h) por países (2017)

- **GDP por hora trabajada.** En la línea de lo anterior, como se puede observar en la figura 24, los países con un GDP por hora trabajada mayor fueron Irlanda (102.39\$), Luxemburgo (98.57\$), Noruega (82.03\$).... En el otro extremo (teniendo en cuenta que no hay datos de países como Camboya, Birmania, Bangladesh...) el país con un GDP por hora trabajada menor es Colombia (13.51\$), seguido de Sud África, México, Costa Rica...

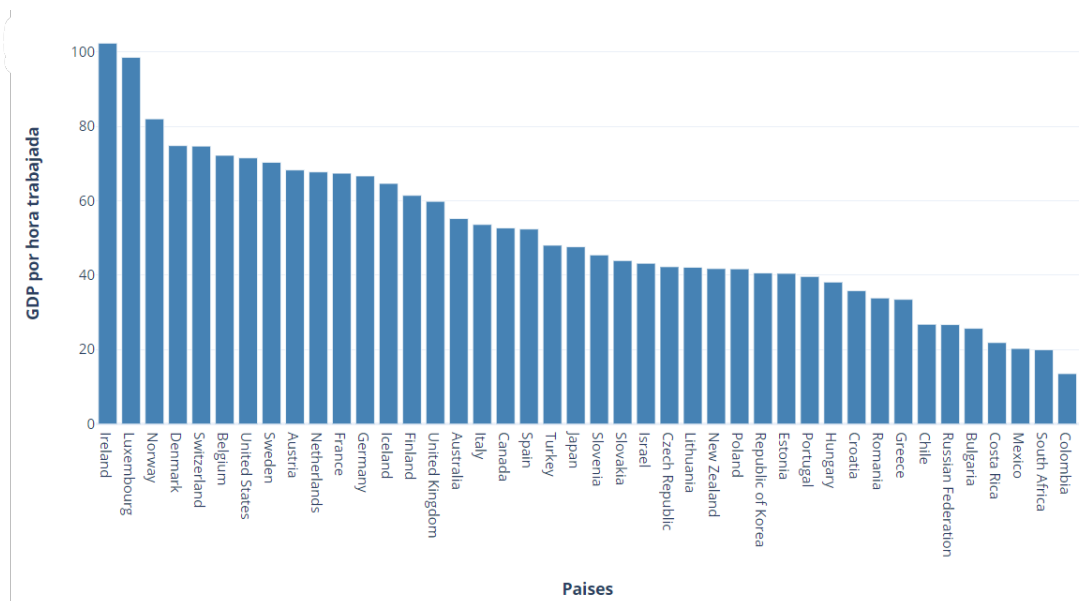


Figura 24: GDP por hora trabajada por países (2019)

- **Salario diario.** En este marco, como se puede observar en la figura 25 , la situación respecto del salario mediano diario sigue la tendencia observada: los países con mayor salario diario son Luxemburgo (63.75\$), Noruega (62.14\$), Suiza (58.08\$), Estados Unidos (57.21\$)... Por otro lado, los países con un menor sueldo diario son Bangladesh (3.74\$), Sudáfrica (4.32\$), Filipinas (4.98\$), Indonesia (5.24\$), Birmania (5.67\$)...

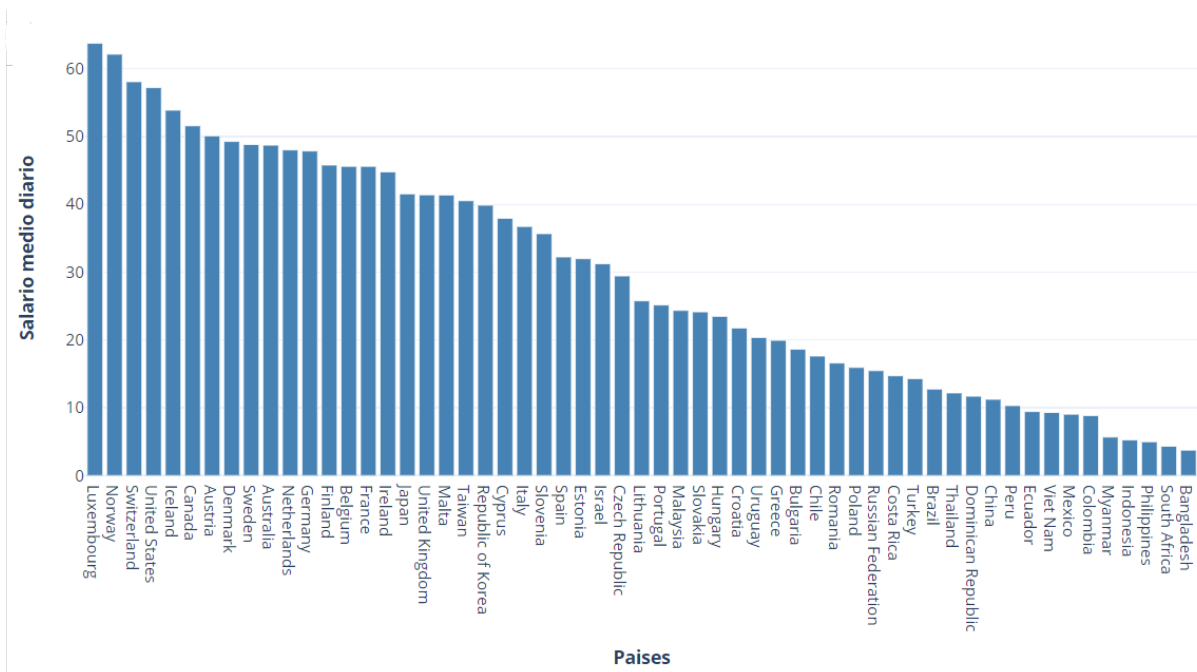


Figura 25: Salario mediano diario por países (2019)

- **Días de vacaciones.** Respecto a los días de vacaciones y festivos nacionales (variable no considerada hasta ahora debido a la poca cantidad de datos que presenta), en la figura 26 se puede observar como Alemania e Italia son los dos países en los que los trabajadores gozan de más días de descanso al año (con 42.5 y 41.5 días respectivamente). Por contra, Canadá y Estados Unidos son los países (con datos) en los que los trabajadores tienen menos días de vacaciones al año, y es que, en el caso de Estados Unidos, según el Foothold America [9], los trabajadores estadounidenses no tienen derecho legal a ninguna fiesta pagada.

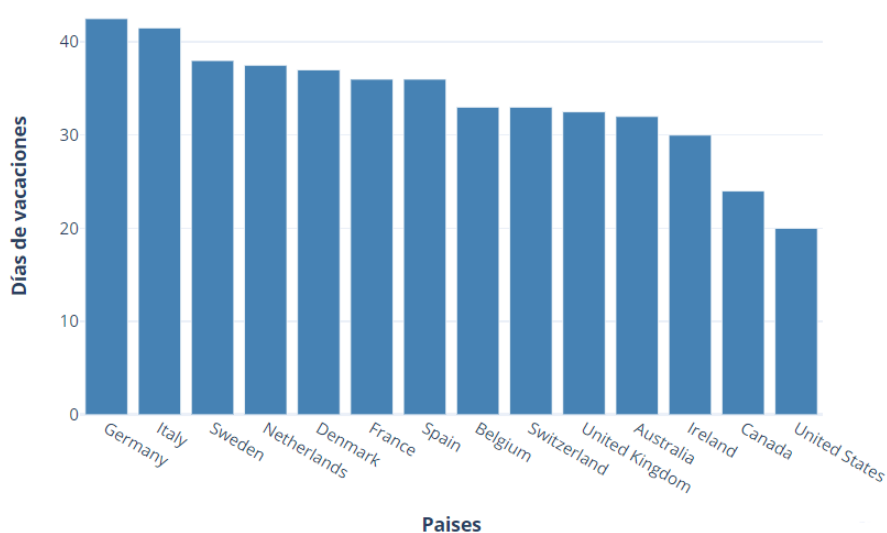


Figura 26: Días de vacaciones al año por países (2000)

3.3.2. ¿Cómo han evolucionado las variables?

Para ver como han evolucionado las variables a lo largo del tiempo, con el objetivo de poder mostrar un análisis claro, en los siguientes gráficos de líneas se han considerado únicamente unos pocos países.

Con esto, como se puede observar en la figura 27, la evolución de cada país respecto de las horas trabajadas es la siguiente: se observa una tendencia a la baja en países como Estados Unidos, Suiza, Alemania, Noruega, Dinamarca... En este último, por ejemplo, en 2019 se trabajó un aproximadamente un 33 % menos que en 1950.

Por contra, países como Grecia, Bangladesh, Camboya o Birmania presentan un tendencia estable y en algunos puntos al alza como es el caso de de Camboya, donde en 2019 se trabajó casi un 13 % más que en 1993 (primer año en el que se tiene un dato en este país).

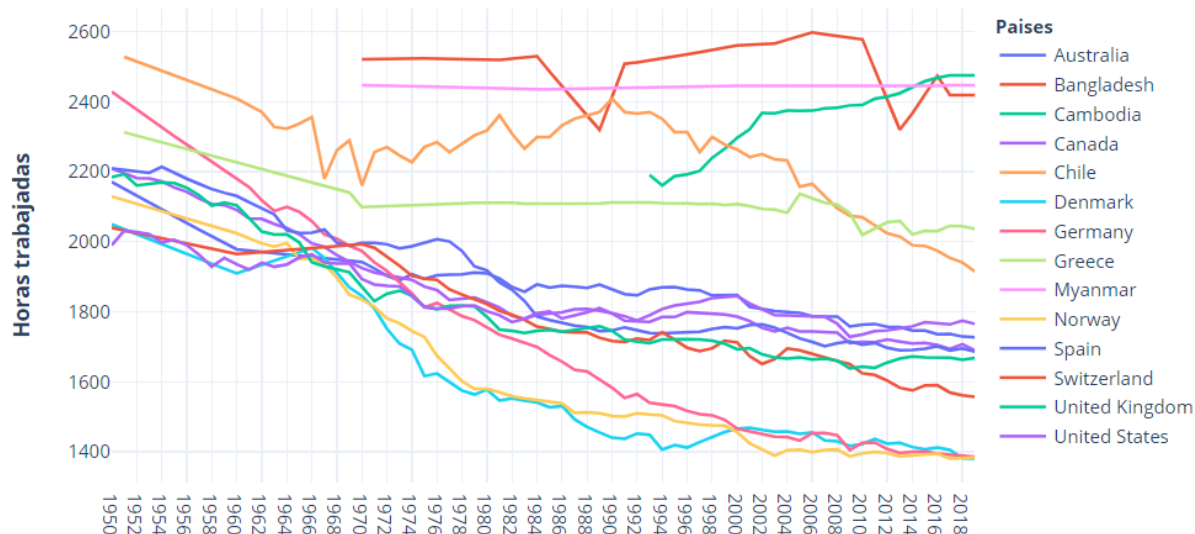


Figura 27: Evolución de las horas anuales trabajadas por países (1950-2019)

Por contra, la situación en el resto de variables analizadas minuciosamente es la siguiente:

- **Productividad.** En la línea de lo anterior, en la figura 28 se observa como, en general los países Europeos y Norteamericanos presentan una clara tendencia al alza (en este grupo destaca especialmente Noruega con una tendencia especialmente irregular pero siendo siempre el país más productivo, salvando como se ha visto antes a Irlanda). Con esto, por ejemplo, los empleados noruegos fueron en 2017 casi un 851 % más productivos que en 1950.

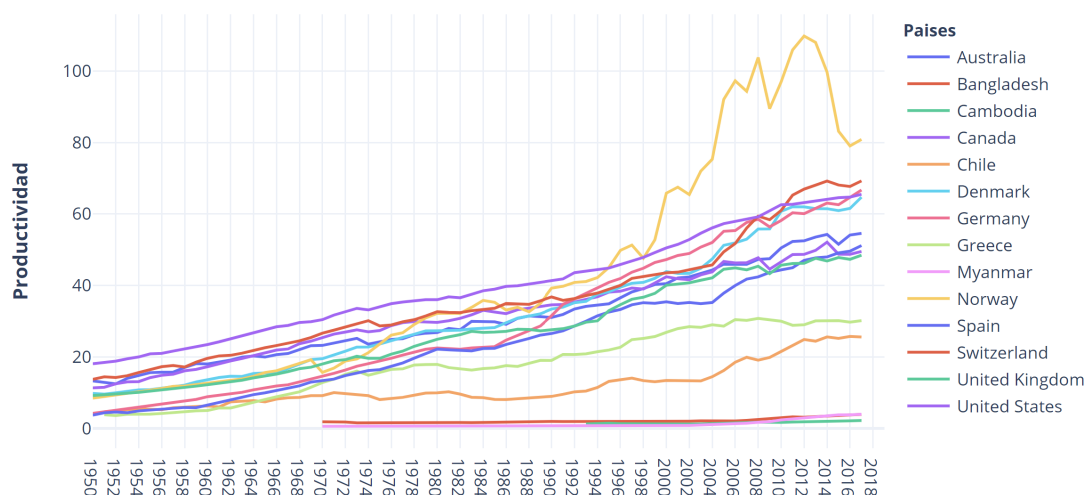


Figura 28: Evolución de la productividad (\$/h) por países (1950-2017)

- **GDP por hora trabajada.** En este caso, como los países asiáticos no presentan datos en esta variable, en la figura 29 se observa como en todos los países utilizados para su realización el GDP por hora trabajada ha aumentado desde 1970 hasta 2019.

Destaca, entre otros países, Suiza cuyos datos en 2019 fueron un 82.62 % mayores que en 1970.

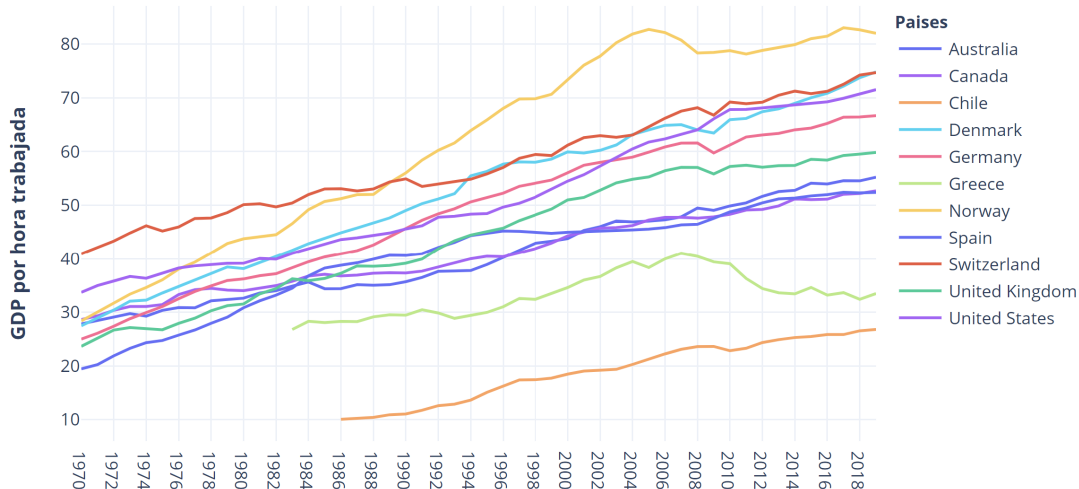


Figura 29: Evolución del GDP por hora trabajada por países (1970-2019)

- **Salario mediano diario.** En la línea de los visto anteriormente, como se puede ver en la figura 30, a excepción de Birmania y Bangladesh (puesto que de Camboya no hay datos en esta variable) cuyo salario diario se ha mantenido ciertamente estable en comparación con el resto de países considerados, pues por ejemplo, en Bangladesh los trabadores cobraron en 2019 un 68.47 % más que en 1981, mientras que, por ejemplo, los datos de 2019 en Noruega fueron un 151.78 % mayores que en 1981.

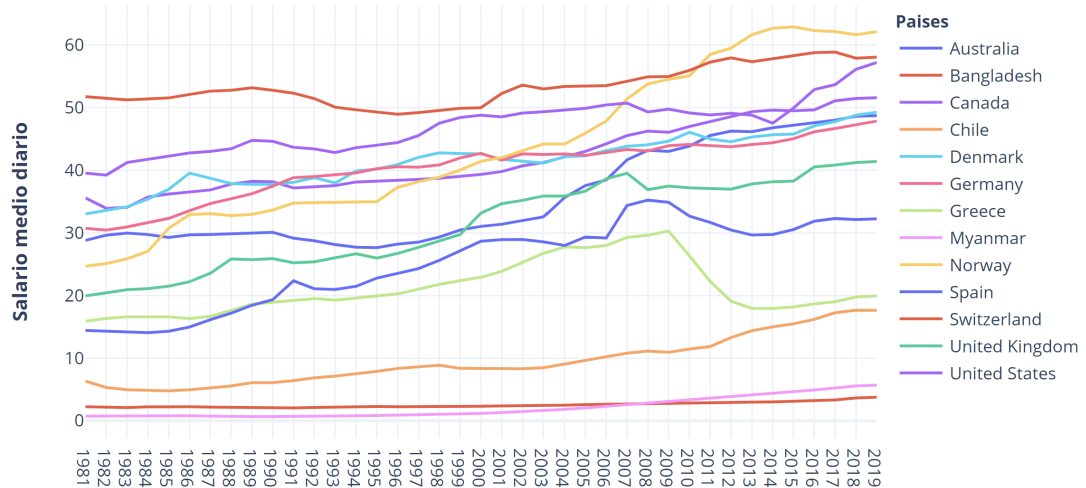


Figura 30: Evolución del salario mediano diario por países (1981-2019)

- **Días de vacaciones y festivos nacionales anuales.** En la figura 31 se puede observar como, en aquellos pocos países donde hay datos de esta variable, los días de vacaciones al año han aumentado a lo largo del tiempo.

Cabe destacar, sin embargo, el caso de EEUU pues en este caso los datos se han mantenido estables con una tendencia a la baja en 1990.

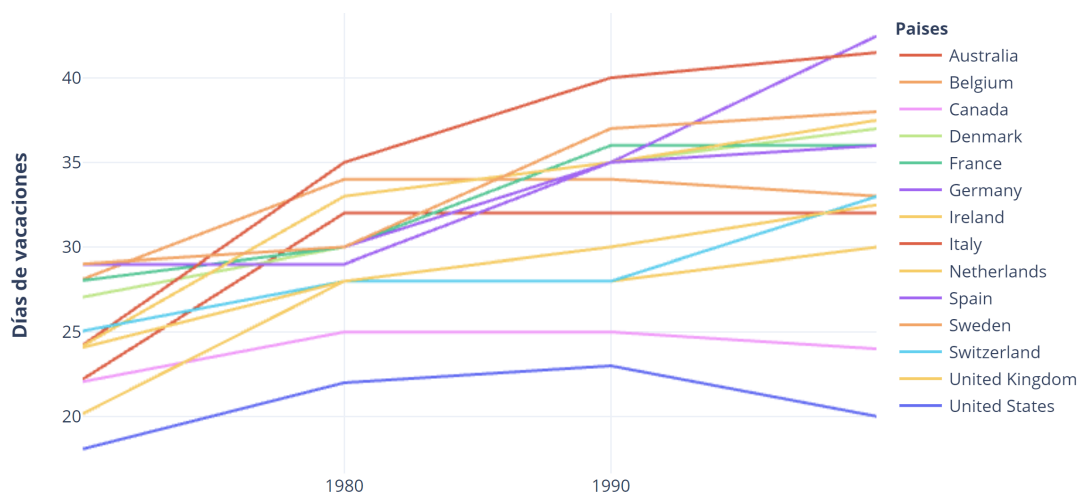


Figura 31: Evolución de los días de vacaciones anuales por países (1950-2000)

3.4. ¿Son los países más ricos los que más trabajan ?

Para contestar a la pregunta de si los países más ricos son los que trabajan más, se ha utilizado un gráfico de correlación que relaciona las horas anuales trabajadas con el GDP (del lado del gasto, aunque el gráfico es muy similar para el lado de la producción) per cápita. Esto puede ser observado en la figura 32.

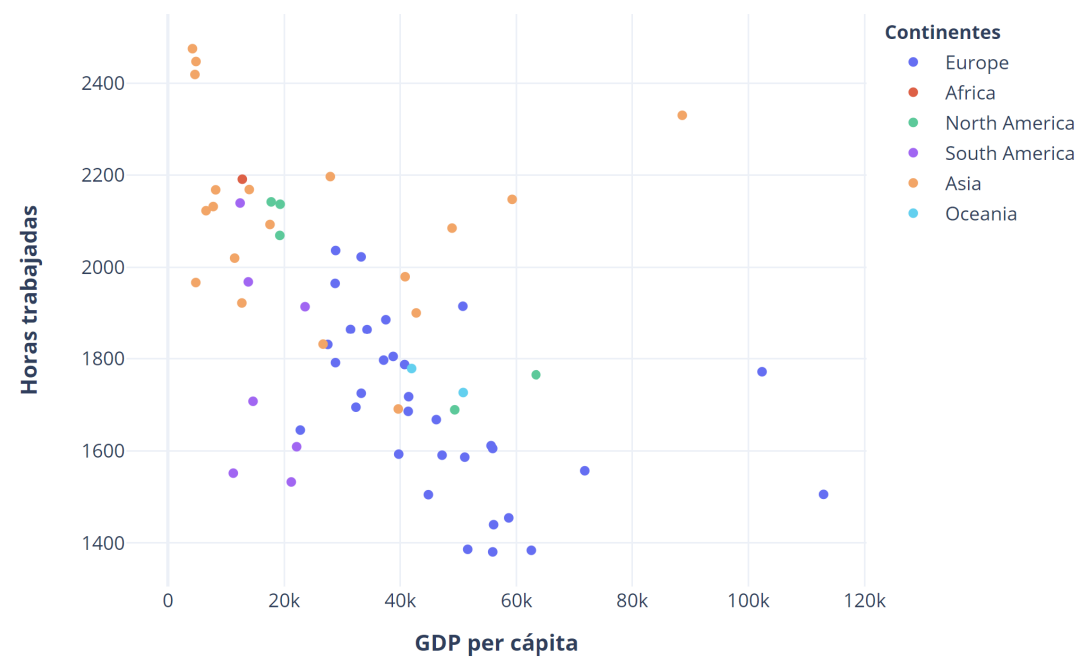


Figura 32: Gráfico de dispersión: Horas anuales trabajadas/GDP del lado del gasto per cápita

Observando el gráfico, y teniendo en cuenta que el coeficiente de correlación lineal entre estas dos variables es de -0.58, se puede concluir que ambas variables tienen una relación inversa, es decir, a medida que se reducen las horas trabajadas aumenta el GDP per cápita, y al contrario. Esto significa que en los países más pobres en realidad tienden a trabajar más, y en ocasiones muchísimo más que en los países más ricos.

Particularmente cabe destacar algunos países que se salen de la línea general: por ejemplo, Singapur presenta un GDP per cápita mucho mayor que otros países que trabajan un número de horas trabajadas como Sud África o Bangladesh. En el lado opuesto, países como Irlanda o Luxemburgo presentan un GDP per cápita muchísimo mayor que otros países que trabajan lo mismo que los inmediatamente citados, como por ejemplo Estados Unidos y Suiza.

3.5. ¿Aumenta la productividad a medida que disminuyen las horas de trabajo?

El factor clave que puede explicar el punto anterior es el crecimiento de la productividad, pues la riqueza de estos países no es debido a que trabajan más horas sino a que el retorno económico por hora de trabajo es mayor (es decir, el número de insumos que se convierten en productos por hora es mayor). Con esto, la productividad tiene el mismo tipo de relación (-0.69) con las horas trabajadas que en el caso anterior, por tanto, a medida que aumenta el número de horas trabajadas disminuye la productividad, tal y como se puede ver en la figura 33.

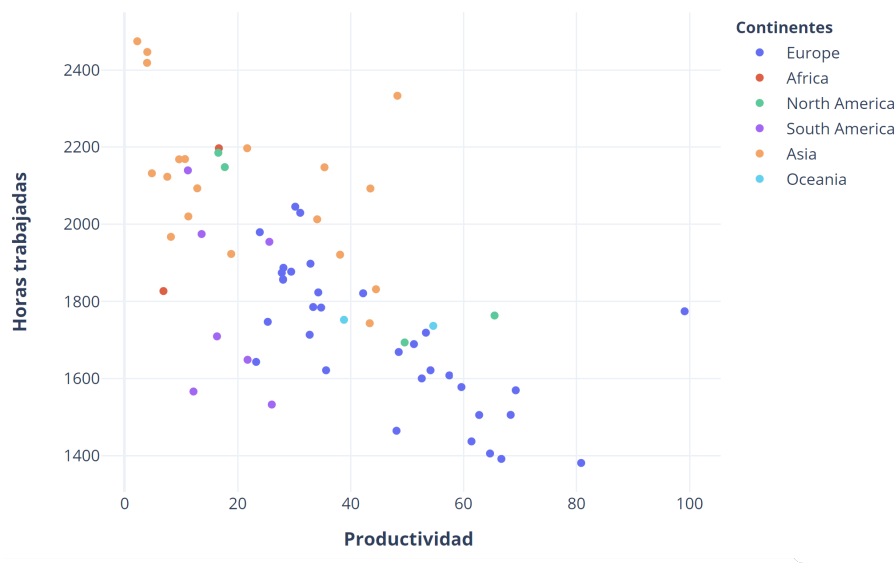


Figura 33: Gráfico de dispersión: Horas anuales trabajadas/Productividad (\$/h)

Sin embargo, de la misma forma que en el caso anterior evidentemente la relación no es perfecta pues, por ejemplo, en Estados Unidos y Singapur, las personas trabajan muchas más horas que países con una productividad similar como Francia o Suiza en el caso de EEUU, y Taiwán o Turquía en el caso de Singapur.

En este punto cabe crear un modelo de regresión lineal simple con el objetivo de poder predecir la productividad media de un trabajador dependiendo de las horas que trabaje, para lo cual es necesario invertir los ejes de forma que las horas trabajadas ahora están situadas en el eje de abscisas y la productividad en el de ordenadas. Este modelo se puede observar en la figura 34 y la estimación de la recta de regresión es la siguiente:

$$-0,0538x + 134,44320746963112,$$

donde x son las horas trabajadas

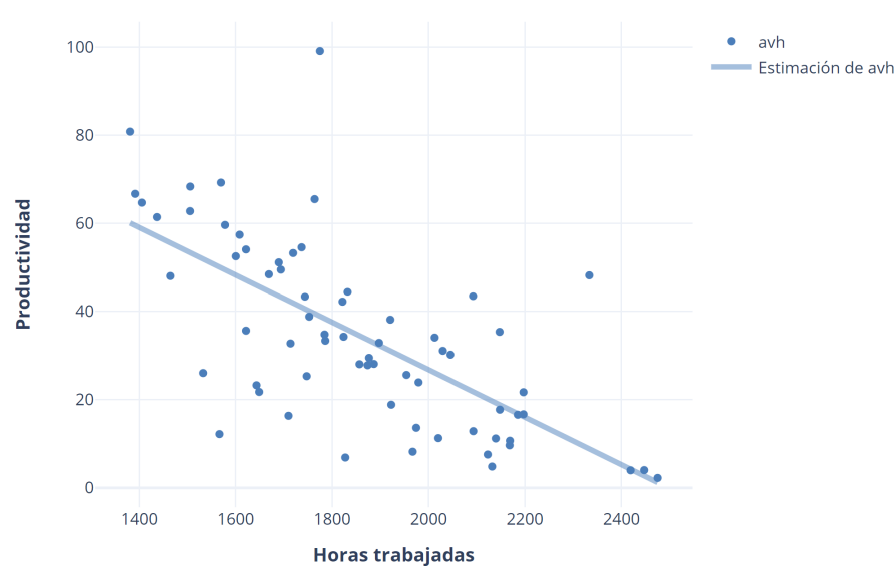


Figura 34: Regresión lineal: Horas anuales trabajadas/Productividad (\$/h)

Así, por ejemplo, una persona que trabaje 2277 horas anuales se estima que tendrá una productividad de 11.93 \$/h, mientras que una que trabaje 1533 horas anuales (29.48 h/semana) generará de forma estimada casi 52\$/h.

3.6. ¿Cobran más aquellas personas que más trabajan?

A priori, y visto desde un punto de vista global, las personas que más trabajan deberían ser aquellas cuyo salario sea mayor, sin embargo como puede ser observado en la figura 35 la situación es bien distinta, pues en realidad las personas que más trabajan son las que menos salario diario presentan, mientras que aquellos en países con menos horas trabajadas los empleados perciben un salario mayor.

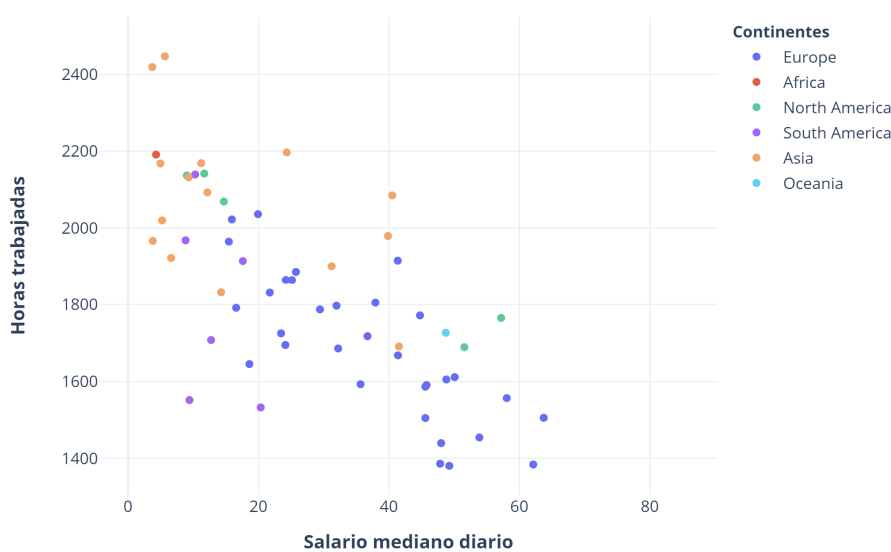


Figura 35: Gráfico de dispersión: Horas anuales trabajadas/Salario mediano diario

3.7. ¿Tienen más calidad de vida aquellas personas que trabajan menos?

Para intentar contestar a esta pregunta se puede analizar la correlación entre las horas trabajadas y variables como el IDH (Índice de Desarrollo Humano) y la puntuación en el World Happiness Report. Esta correlación se puede observar en las figuras 36 y 37.

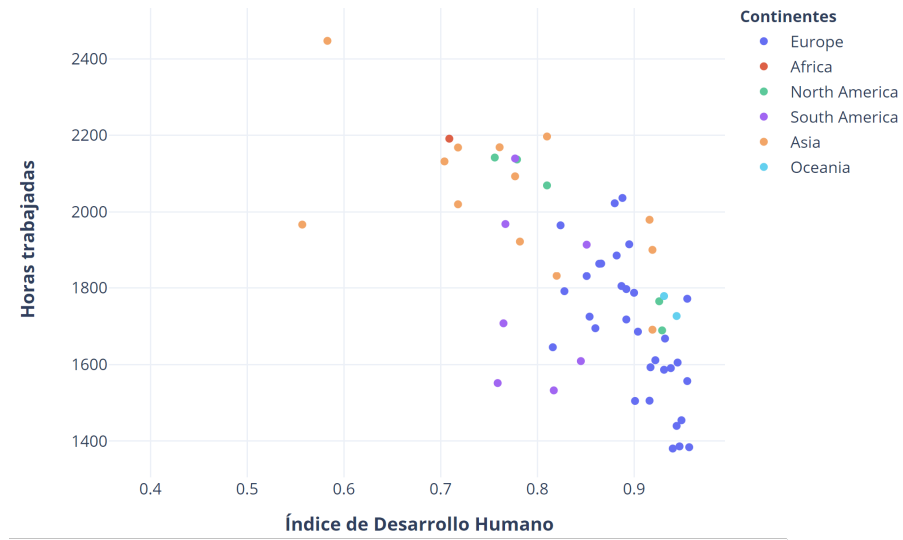


Figura 36: Gráfico de dispersión: Horas anuales trabajadas/Índice de Desarrollo Humano

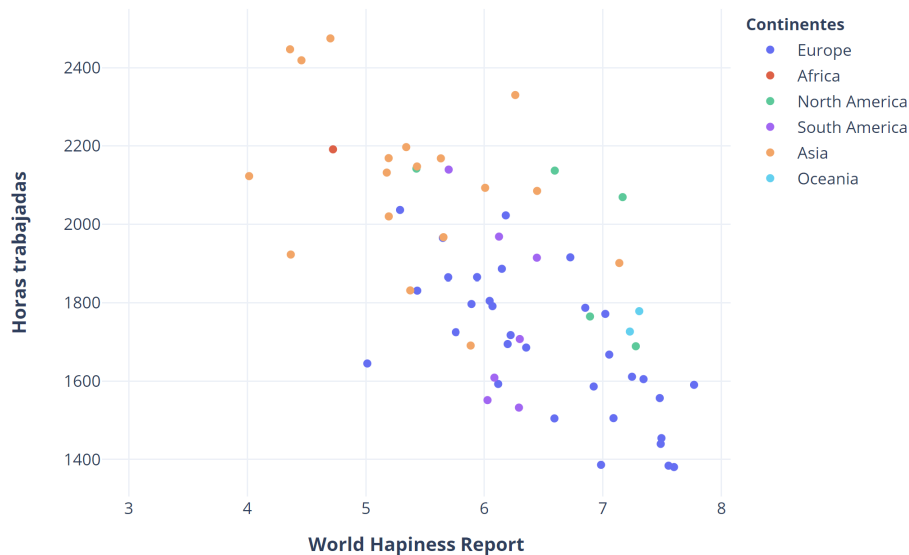


Figura 37: Gráfico de dispersión: Horas anuales trabajadas/Puntuación en el WHR

Como se puede observar en ambas figuras la relación que siguen estas dos variables es inversa (de -0.59 en el caso de IDH y -0.62 en el caso de la puntuación en el World Happiness Report). Por tanto, se puede intuir como a medida que disminuyen las horas de trabajo aumenta la calidad de vida y por tanto, en parte, la felicidad de las personas.

3.8. Modelo predictivo

3.8.1. Predicción global

En la figura 38 se muestra un gráfico en el cual se puede observar la evolución de las horas trabajadas en el mundo (con la línea delgada), y además una predicción generada mediante regresión simple para los próximos 5 años (desde el 2022) (con la línea más gruesa) utilizando la función cuadrática

$$0,02313x^2 - 96,8657x + 103106,9347$$

donde x es el año del cual se pretende saber el número de horas trabajadas.

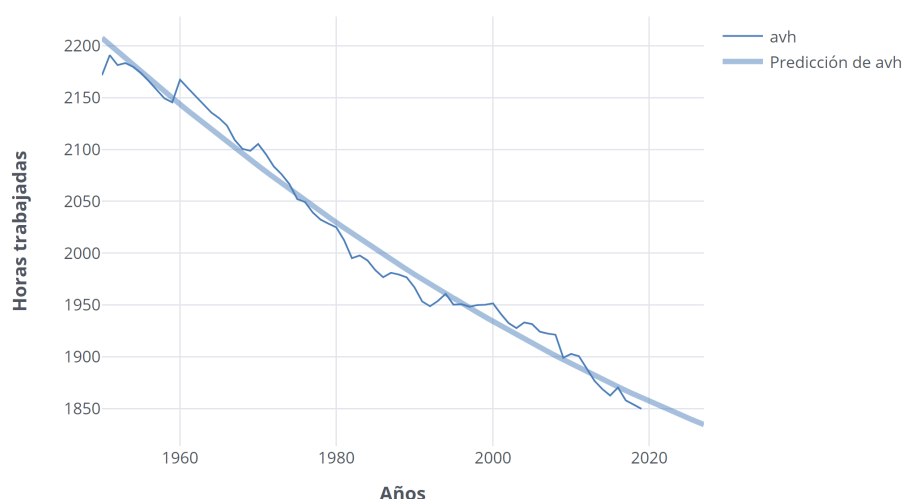


Figura 38: Regresión lineal: horas anuales trabajadas

Mediante el modelo de regresión es posible estimar que, por ejemplo, en 2070 se trabajarán 32.78 horas a la semana.

3.8.2. Predicción en España

Antes de analizar cual será la situación del estado español en un futuro, es importante analizar cual ha sido su evolución a lo largo del tiempo. En la figura 39, se puede ver una tendencia claramente bajista, pero que se pronunció todavía más a partir del año 1975, año en el que acabó la dictadura franquista y comenzó la transición lo que supuso la integración con países europeos entrando, por ejemplo, en la CEE (Comunidad Económica Europea). Con esto, en España en el 2019 se trabajó un 23.72 % menos que en el año 1950.

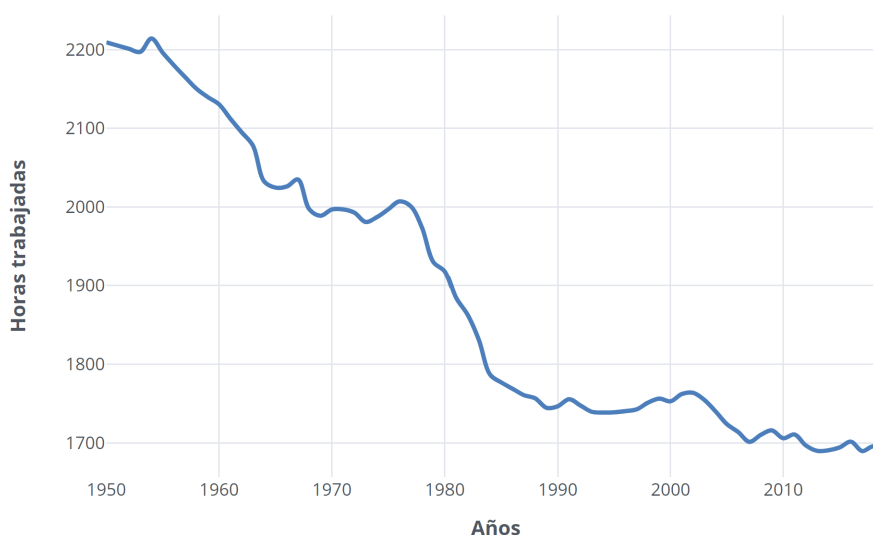


Figura 39: Evolución de las horas trabajadas en España (1950-2019)

En este contexto, es interesante conocer cual será la situación en un futuro para lo cual se utilizará la misma herramienta que para el análisis global. En la figura 40 se puede ver como, en este caso, y teniendo en cuenta que es un modelo algo limitado, en España en

2027 se trabajarán 1687 horas al año, lo que indica que las horas trabajadas aumentarían un 0.11 % respecto a 2019.

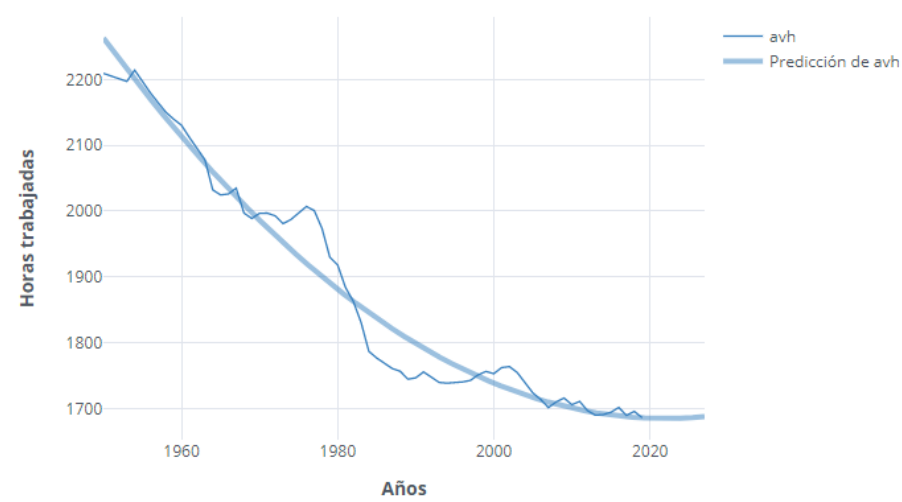


Figura 40: Regresión lineal: horas trabajadas en España

La curva de regresión en este caso es la siguiente:

$$0,1117x^2 - 451,7068x + 458348,4908$$

4. Interpretación de los datos

A partir de los datos obtenidos podemos concluir que, en general, las horas trabajadas se han reducido considerablemente entre 1950 y 2019, mientras que la productividad, GDP por hora trabajada y el salario diario han incrementado en este mismo periodo de tiempo. Sin embargo, cabe destacar que la situación es tremendamente dispar entre los diferentes países y, por consiguiente, entre los diferentes continentes:

Cabe destacar, por ejemplo, como los países europeos son, junto con los de Oceanía, EE.UU. y Canadá, los lugares en los que menos se trabaja. Por contra, los países donde más se trabaja son los asiáticos y los africanos.

Por otra parte, los países más productivos, con más GDP por hora trabajada y con más salario, son los países europeos y de Oceanía; y aquellos cuyos datos de estas variables son menores son los asiáticos y africanos.

Esta diferencia puede ser debida, entre otros factores, a que las empresas tienen sus sedes sociales y tributarias en los países del norte global, de forma que profesiones del sector servicio (diseño, dirección...) están situadas en estos países. Por contra, las actividades de producción se deslocalizan a países con una regulación ambiental y laboral deficiente, donde las condiciones laborales de los trabajadores (horas trabajadas, salario...) son extremadamente cuestionables.

5. Estado del arte

Se puede introducir el estado del arte como la última situación de la que se tiene constancia de una temática concreta (en este caso, las horas trabajadas), dando a conocer de esta forma los conocimientos y resultados previos al estudio actual y generando nuevas hipótesis de investigación para dar comienzo a cualquier estudio.

Así pues, aplicando dicho concepto del estado del arte a nuestro proyecto, podemos situar el cumplimiento de los objetivos generales propuestos de forma que se tengan en cuenta los avances de nuestro trabajo. Por una parte, se dan aquellos objetivos generales en los cuáles nos hemos centrado en conseguir con la máxima información posible, como puede ser el caso de relacionar las horas de trabajo con el GPD de los países, para luego poder lograr acercarnos al cumplimiento de algunos objetivos específicos impuestos como, por ejemplo, comparar la productividad de España con la del resto de países europeos (análisis que hemos evaluado de forma óptima). Por otra parte, hay varios objetivos específicos (SMART) que, aunque hayamos conseguido cumplir los objetivos generales relacionados o de los que derivan, no hemos logrado cumplirlos. Algunos de estos objetivos eran analizar la relación entre la maternidad/paternidad y la productividad y el de abordar la jornada laboral de 4 días, debido a la dudosa relación con los objetivos y que no se llegaba a contestar del todo a nuestras preguntas.

A lo largo del proyecto, hemos encontrado diversos estudios y noticias con la que comparte temática nuestro trabajo. La mayoría de esos estudios han encontrado cierta correlación entre las horas trabajadas y la productividad [10], afirmando que una reducción de las horas genera una mayor productividad [11]. Confirman que esto se debe a los impactos negativos que generan las "largas horas de trabajo", y que afectan a la salud de los trabajadores [12], implicando una reducción de la productividad [13]. La reducción de las horas de trabajo da lugar a una menor sensación de fatiga y agobio, mejorando la predisposición de los trabajadores. Además, varios estudios afirman que lo ideal para lograr una productividad máxima es una jornada laboral de 48 horas a la semana [14]. También se trata la reducción de horas de trabajo a lo largo de los años y el teletrabajo que ha traído la pandemia estos últimos años [15].

En la línea de lo explicado, también se muestran diferencias entre países cercanos [16]. Se afirma que en Europa, el sector que más sufre de largas jornadas es el sector de servicios, especialmente en países como España.[10]

Ahora bien, todos estos estudios tienen cierta carencia: el futuro. Ninguno de ellos va más allá del presente, solamente hacen análisis de la actualidad, sin una visión que pueda decirnos como será nuestra vida laboral, o como apunta que será. Este proyecto (además de confirmar todo lo anteriormente estudiado) hace una predicción realista y basada en los datos de los que se disponen, así como una mayor cercanía al usuario, que podrá comparar datos presentes y pasados mediante las gráficas interactivas que se encuentran en nuestra página web.

6. Página Web

La página web con la que cuenta nuestro proyecto ha sido realizada con la herramienta "WebNode", y su dirección URL es la siguiente: <https://horastrabajadas.webnode.es/>. Esta cuenta con cinco secciones:

- **Inicio.** Una entrada que explica brevemente qué es lo que ofrece nuestra web: informar sobre las diferencias sobre las horas trabajadas entre distintos países, las relaciones entre distintas variables mediante gráficos y un modelo predictivo para ver cómo evolucionará el mercado laboral.
- **Contexto estadístico.** En esta entrada, se habla sobre la obtención de los datos, la selección de los datos y los análisis previos, la preparación de los datos (faltantes y correlación entre las distintas variables) así como más análisis con gráficos de dispersión.
- **Análisis:**
 1. **Análisis global.** Permite observar y analizar mediante gráficos la evolución de las horas de trabajo anuales, la evolución de la productividad, el GPD por hora trabajada y la evolución del salario medio, todo a nivel global.
 2. **Análisis por continentes.** Este apartado está más enfocado a nivel continental, poniendo primero en situación a quien lo lee, para después ofrecerle un gran número de análisis y la evolución de las distintas variables.
 3. **Análisis por países.** En este apartado nuevamente se pone en contexto al visitante de la página, para después ofrecerle gran cantidad de análisis y la evolución de las variables.
 4. **Predicción.** En este apartado, se hace una predicción a partir de los datos de los que hemos dispuesto a nivel global y nacional(España) sobre las horas que se trabajarán de aquí a unos años.
- **Realiza tu análisis.** En este apartado se ubica el enlace a una página en la cual el usuario puede diseñar sus propios análisis a partir de nuestra base de datos, pudiendo variar años, países y variables. El link de dicha página es el siguiente <https://horas-trabajadas.herokuapp.com/>
- **Sobre nosotros.** En este apartado, se habla sobre el equipo. Aquí está la misma información de la que se ha dispuesto en este mismo documento en el apartado «Equipo» en Introducción.

7. Documentos y entregables

Como documentos y entregables de este proyecto, se ofrecerán los siguientes:

- Un documento en formato .pdf realizado mediante Overleaf (L^AT_EX)
- Una presentación como documento gráfico, que será utilizada como defensa de este proyecto
- Una página web
- Las bases de datos utilizadas y modificadas para el proyecto.
- El archivo .ipynb con el código Python utilizado para general los análisis

8. Desarrollo de la EPT

8.1. Conclusiones

Las horas trabajadas, en general, han disminuido, mientras que la productividad, el GDP por hora trabajada y el salario diario han aumentado. Además, se concluye que en Europa, Oceanía y Norte América se trabaja menos que en África y Asia, mientras que ocurre lo contrario con las otras variables comentadas.

Con esto, se concluye que existe una relación (inversa) considerable entre el número de horas trabajadas y variables como el GDP per cápita, el IDH, la puntuación en el WHR, la productividad y el salario diario.

Comentar finalmente que no se han podido tratar cuestiones como el análisis de la paternidad/maternidad en la productividad del trabajo debido a la falta de datos al respecto.

8.2. Propuesta de futuro

Una interesante manera de continuar este proyecto sería a través de todo lo acontecido durante y después de la pandemia de COVID-19. Actualmente, no se contaba con suficientes datos (los cuales carecían en gran medida de precisión) para realizar un análisis de la situación. Habrá que esperar un tiempo antes de contar con unos datos óptimos en lo referido al teletrabajo y la crisis laboral que supuso la cuarentena y la posterior situación de "nueva normalidad" para el mercado laboral y las horas que realizaban los trabajadores. Otra manera de continuar el trabajo es orientarlo hacia el estudio de la jornada laboral de 4 días, ya que es un campo en el que no hemos llegado a profundizar y que, sin embargo, podría sernos útil en un futuro para determinar si ello sería algo viable.

9. Bibliografía

- [1] Feenstra, Robert C., Inklaar R. I& Timmer M.P. (2015). The Next Generation of the Penn World Table. American Economic www.ggd.net/pwt
- [2] Young, W. (2022, May 7). How Many Workweeks in a Year?. Employment Security Commission. <https://www.ncesc.com/how-many-workweeks-in-a-year/#:~:text=0n%20average%2C%20Americans%20have%2048,the%20majority%20of%20the%20world>
- [3] Productivity - GDP per hour worked - OECD Data. (2021). theOECD. <https://data.oecd.org/lprdt/gdp-per-hour-worked.htm>
- [4] Daily median income. (2013). Our World in Data.https://ourworldindata.org/grapher/daily-median-income?country=OWID_WRL%7EESP%7EKOR%7EMDG
- [5] Huberman M. I& Minns C. (2007). The times they are not changin': Days and hours of work in Old and New Worlds, 1870–2000. Exploration in Economic History.<https://www.sciencedirect.com/science/article/abs/pii/S0014498307000058>
- [6] Human Development Reports. (2020). United Nations Development Programme<https://hdr.undp.org/en/indicators/137506#>
- [7] World Happiness Report. (2019, November 27). Kaggle.<https://www.kaggle.com/datasets/unsdsn/world-happiness>
- [8] Amadeo, K. (2022, January 28). When and Why Did the Stock Market Crash in 2008? The Balance. Retrieved May 15, 2022, from <https://www.thebalance.com/stock-market-crash-of-2008-3305535>
- [9] Foothold America. (2022, May 13). What is the US Paid Holiday Entitlement? — Foothold America. Foothold America — US Business Expansion Experts. Retrieved May 18, 2022, from <https://www.footholdamerica.com/faqs/how-many-paid-holiday-days/#:~:text=US%20workers%20are%20not%20legally,of%20paid%20holiday%20each%20year>
- [10] Collewet M. I& Sauermann J. (April 2017). Working Hours and Productivity. IZA, Institute of Labor Economics. <https://ftp.iza.org/dp10722.pdf>
- [11] Kodz J., Davis S., Lain D., Strebler M., Rick J., Bates P., Cummings J. I& Meager N. (October 2003). Employment Relations Research Series. Department of Trade and Industry. <https://www.employment-studies.co.uk/report-summaries/report-summary-working-long-hours-review-evidence-volume-1-%E2%80%9393-main-report>
- [12] Wong K., Chan A. I& Ngan S. C. (2019). The Effect of Long Working Hours and Overtime on Occupational Health: A Meta-Analysis of Evidence from 1998 to 2018. International journal of environmental research and public health.<https://doi.org/10.3390/ijerph16122102>
- [13] Luxton E. (March 4, 2016). Does working fewer hours make you more productive?. We Forum (org). <https://www.weforum.org/agenda/2016/03/does-working-fewer-hours-make-you-more-productive/>
- [14] Luxton, E. (2020, February 5). Does working fewer hours make you more productive? World Economic Forum. <https://www.weforum.org/agenda/2016/03/does-working-fewer-hours-make-you-more-productive/>

- [15] Yang L., Holtz D., Jaffe S. (September 15 2021). Home working brings longer hours, fewer phone calls. *Nature Human Behaviour*. <https://rdcu.be/cNvSq>
- [16] Bick A., Brüggemann B. I& Fuchs-Schündeln N. (February 22, 2019). Hours Worked in Europe and the United States: New Data, New Answers. *The Scandinavian Journal of Economics*. <https://doi.org/10.1111/sjoe.12344>