



# Micropráctica de Evaluación

Diseño de Sistemas Interactivos

Curso 2018/2019



100363803

Denys Porynets

100346190

Pablo Cañas Castellanos

100346220

Alejandra Coca Martín

100329967

Julio Albertos León

# ÍNDICE

1. Experimento con usuarios	2
2. Análisis	7
3. Conclusiones	15

# 1. Experimento con usuarios

## Objetivos

En este experimento, se quiere evaluar la utilidad, usabilidad, eficacia, eficiencia y experiencia del usuario del software Microsoft Powerpoint, diseñado para componer presentaciones (sucesión de diapositivas).

Se van a tener en cuenta los siguientes requisitos de usabilidad:

1. La aplicación tiene que ser fácil de usar aún para usuarios principiantes.
2. La aplicación tiene que proporcionar una navegación intuitiva.
3. La interfaz y el resultado final tiene que ser visualmente agradable.
4. El sistema tiene que facilitar el trabajo repetitivo.
5. El sistema tiene que ser percibido como útil.

## Participantes

El perfil de los participantes del experimento se expone a continuación:

- En cuanto a demografía, buscamos un grupo de participantes muy heterogéneo: rango de edad entre 15 a 60 años, ambos sexos, de cualquier localidad española y de condición económica y social diversa.
- Sin embargo, buscamos participantes que estén familiarizados con el uso de ordenadores y con un nivel básico de ofimática. No queremos expertos en el manejo de ordenadores, pues muy probablemente estos usuarios no tengan dificultad con el software, pero queremos un mínimo de nivel de habilidad.

Como queremos obtener un conjunto de datos lo suficientemente amplio, hemos estimado el número de participantes en 500. Aunque los costes de la evaluación serán mayores, pensamos que este número nos aportará estadísticas veraces sobre la eficiencia del sistema.

## Tareas

El escenario que vamos a evaluar para este experimento es el siguiente: **Creación de una presentación con 5 diapositivas.**

Algunos de los requisitos del escenario son los siguientes:

- La primera será una carátula.
- El resto, tendrá un logo arriba a la izquierda, y abajo a la derecha el número de diapositiva.
- Cada diapositiva, a partir de la primera, tendrá un título, siempre con el mismo estilo y tipo de letra, y en el cuerpo de la diapositiva tendrá un esquema con 5 puntos, precedidos de un círculo azul.

- El diseño de todos los elementos usados deberá ser consistente para todas las diapositivas.
- Usar todos los recursos posibles de la herramienta para trabajar lo menos posible (de la forma más eficiente posible) desarrollando la presentación.

Con estas tareas, se busca obtener medidas de eficiencia para los cinco requisitos de usabilidad que hemos definido en nuestros objetivos.

## Indicadores

A continuación exponemos los parámetros que se van a medir para evaluar la eficiencia del sistema:

1. Tiempo tomado para llevar a cabo la tarea, en segundos.
2. Número de clicks realizados para llevar a cabo la tarea.
3. Número de errores cometidos para llevar a cabo la tarea.
4. Número de 'Deshacer' realizados para llevar a cabo la tarea.

Pensamos que estos cuatro indicadores son suficientes para realizar un análisis cuantitativo razonable.

## Guías experimentales

Documento para los usuarios:

Estimado usuario,

Muchas gracias por participar en nuestro experimento. En dicho experimento, le propondremos una tarea a realizar con el software Powerpoint, para evaluar la eficiencia del mismo y poder ofrecer mejores productos y servicios a nuestros clientes.

Para ello, se le proporcionará con el equipamiento necesario y se le pedirá que complete una tarea con la herramienta. Es importante denotar que no hay límite de tiempo y que no se están evaluando sus habilidades, sino validando la eficacia del software para situaciones posiblemente desconocidas para un usuario.

La tarea a ejecutar es la siguiente: **Creación de una presentación con 5 diapositivas.**

Estas diapositivas deben cumplir los siguientes requisitos:

- La primera será una carátula.
- El resto, tendrá un logo arriba a la izquierda, y abajo a la derecha el número de diapositiva.

- Cada diapositiva, a partir de la primera, tendrá un título, siempre con el mismo estilo y tipo de letra, y en el cuerpo de la diapositiva tendrá un esquema con 5 puntos, precedidos de un círculo azul.
- El diseño de todos los elementos usados deberá ser consistente para todas las diapositivas.
- Usar todos los recursos posibles de la herramienta para trabajar lo menos posible (de la forma más eficiente posible) desarrollando la presentación.

El experimento será grabado por una cámara. Sin embargo, es importante remarcar que la cámara apuntará en todo momento al dispositivo, por lo que su anonimato quedará intacto en todo momento.

También es importante explicar que no vamos a almacenar ningún dato de carácter personal, como el nombre o la dirección. Sin embargo, por la Ley Orgánica de Protección de Datos, necesitamos que dé su consentimiento para que almacenamos datos demográficos para fines estadísticos, como edad, sexo o localidad.

Por favor, firme debajo para expresar su consentimiento. Nuestro asistente le facilitará más instrucciones y le resolverá cualquier otra duda.

Firma:

Documento para los observadores:

El usuario va a proceder a realizar un experimento con el software Powerpoint, para evaluar la eficiencia del mismo y poder ofrecer mejores productos y servicios a nuestros clientes.

Para ello, se le proporcionará con el equipamiento necesario y se le pedirá que complete una tarea con la herramienta. Es importante denotar que no hay límite de tiempo y que no se están evaluando sus habilidades, sino validando la eficacia del software para situaciones posiblemente desconocidas para un usuario.

La tarea a ejecutar es la siguiente: **Creación de una presentación con 5 diapositivas.**

Estas diapositivas deben cumplir los siguientes requisitos:

- La primera será una carátula.
- El resto, tendrá un logo arriba a la izquierda, y abajo a la derecha el número de diapositiva.
- Cada diapositiva, a partir de la primera, tendrá un título, siempre con el mismo estilo y tipo de letra, y en el cuerpo de la diapositiva tendrá un esquema con 5 puntos, precedidos de un círculo azul.

- El diseño de todos los elementos usados deberá ser consistente para todas las diapositivas.
- Usar todos los recursos posibles de la herramienta para trabajar lo menos posible (de la forma más eficiente posible) desarrollando la presentación.

Aunque el experimento se grabará y se podrá revisar en un futuro, es importante que tome las siguientes medidas cuantitativas para la evaluación del experimento:

1. Tiempo tomado para llevar a cabo la tarea, en segundos.
2. Número de clicks realizados para llevar a cabo la tarea.
3. Número de errores cometidos para llevar a cabo la tarea.
4. Número de 'Deshacer' realizados para llevar a cabo la tarea.

Si el usuario tiene alguna otra duda de carácter administrativo o técnico, se la puede resolver. Sin embargo, no podrá responder ni facilitar guía en la tarea que se le ha encomendado al usuario.

Cuando el usuario le entregue su hoja firmada y así lo exprese, puede comenzar el experimento.

## Cuestionario

En relación al cuestionario de evaluación, hemos elegido el cuestionario SUS (Escala de Usabilidad del Sistema). En este caso, pensamos que es el cuestionario más apropiado puesto que es sencillo y rápido de responder, las preguntas son muy claras y concisas y, por otro lado, es un cuestionario validado por expertos que engloba todos los requisitos que queremos evaluar, convirtiéndose en una técnica madura y confiable para la extracción de resultados.

El cuestionario consiste de las siguientes preguntas:

P1 - Creo que me gustaría utilizar este sistema frecuentemente	1	2	3	4	5
P2 - Encuentro este sistema innecesariamente complejo	1	2	3	4	5
P3 - Pienso que el sistema es fácil de usar	1	2	3	4	5
P4 - Creo que necesitaría soporte técnico para hacer uso del sistema	1	2	3	4	5
P5 - Encuentro las diversas funciones del sistema bastante bien integradas	1	2	3	4	5
P6 - He encontrado demasiada inconsistencia en este sistema	1	2	3	4	5
P7 - Creo que la mayoría de la gente aprendería a hacer uso de este sistema rápidamente	1	2	3	4	5

P8 - He encontrado el sistema bastante incómodo para usar	1	2	3	4	5
P9 - Me he sentido muy seguro haciendo uso del sistema	1	2	3	4	5
P10 - Necesitaría aprender un montón de cosas antes de poder manejar el sistema	1	2	3	4	5

El usuario puntúa cada pregunta de 1 (totalmente en desacuerdo) a 5 (totalmente de acuerdo). Si la respuesta ocupa una posición impar (pregunta de carácter positivo), el sistema escala dicha puntuación directamente al rango 0 a 4. Si la respuesta ocupa una posición par (pregunta de carácter negativo), el sistema escala de 0 a 4, restando previamente a 5 la respuesta del usuario. Esta puntuación, de 0 a 40, se multiplica por 2.5, para así tener una escala de evaluación de 0 a 100.

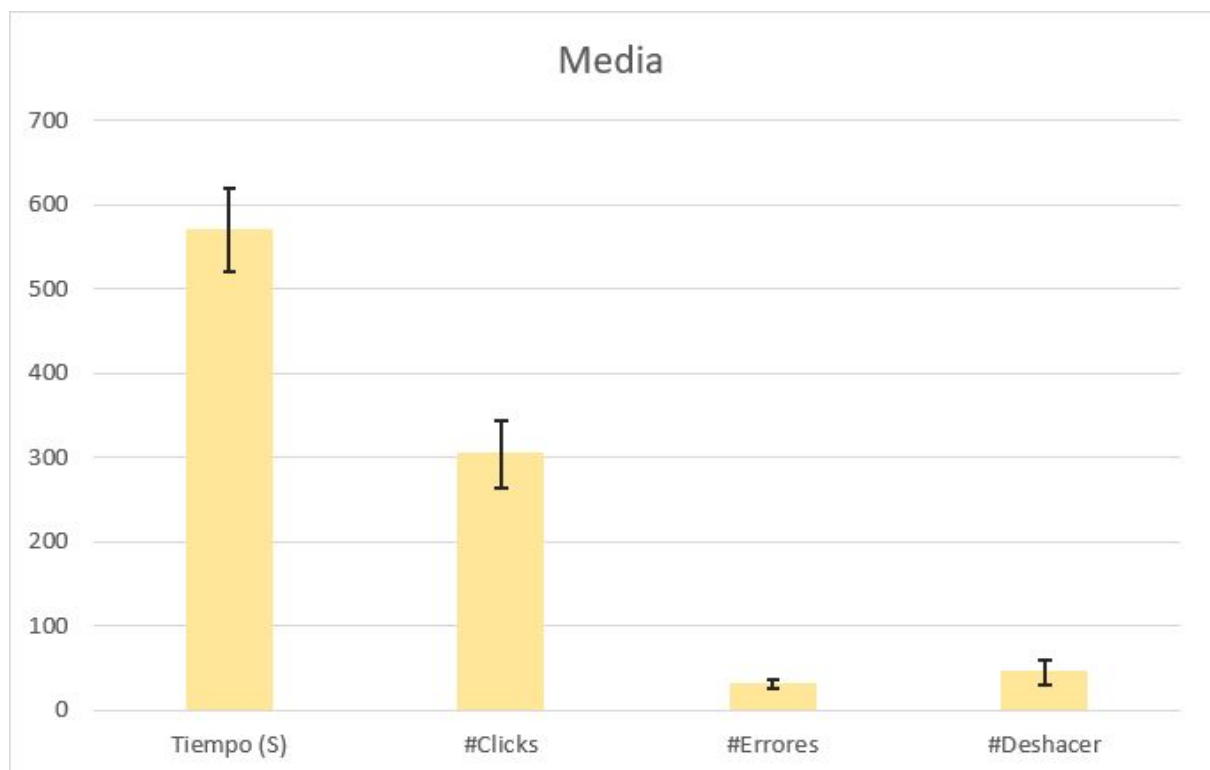
## 2. Análisis

### Análisis cuantitativo

Para nuestro análisis cuantitativo hemos medido los parámetros del apartado de indicadores y hemos obtenido los siguientes datos para los diferentes indicadores obteniendo los estadísticos de promedio y desviación estándar:

Indicador	Media	Desviación Estándar
Tiempo (S)	570	50
#Clicks	304	40
#Errores	32	5
#Deshacer	45	15

El gráfico resultante es el siguiente:



De estos resultados podemos deducir que tiempo medio en completar esta tarea es de 10 minutos, y que no varía mucho entre usuarios de distinto nivel o características. En general, los resultados son bastante homogéneos entre los usuarios, con poca desviación estándar. En el caso de los errores, prácticamente se cometen los mismos. Sin embargo, sí que existe



una diferencia entre el número de vueltas atrás que se realizan, siendo un importante elemento a analizar.

### Análisis del cuestionario

Tras finalizar los cuestionarios para los 500 participantes en la encuesta, obtuvimos los resultados disponibles en el archivo *cuestionariosus.xlsx*. Con esta información procedemos a analizar los resultados de la evaluación de nuestro sistema. A continuación mostraremos algunas de las instancias de nuestra tabla, pero no la tabla completa ya que ésta tiene más de 500 filas:

P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
3	2	4	3	4	3	4	3	4	3
5	1	5	1	5	1	5	1	5	1
4	1	5	2	3	2	5	2	3	2
3	4	3	4	4	3	4	3	4	3
4	2	4	2	4	2	4	2	4	2
4	2	3	2	4	2	3	2	3	1
4	4	3	4	4	3	4	3	2	2
3	3	1	3	4	3	3	3	5	2
3	4	3	3	3	4	3	4	2	3
3	4	3	2	4	3	3	2	3	2
4	2	4	2	4	2	4	2	4	2
4	2	3	2	4	2	3	2	3	1
5	1	5	1	5	2	4	2	5	1
5	1	5	1	5	1	5	1	5	1
4	1	5	2	3	2	5	2	3	2
4	2	3	2	4	2	3	2	3	1
3	2	4	3	2	3	3	2	4	3
4	2	4	2	4	2	4	2	4	2

5	1	5	1	5	1	5	1	5	1
---	---	---	---	---	---	---	---	---	---

En esta tabla podemos ver como la primera fila indica qué columna contiene la información de cada una de las 10 preguntas del cuestionario y cada fila es una instancia de un usuario participando en él y dándole un valor del 1-5 a cada una de las preguntas.

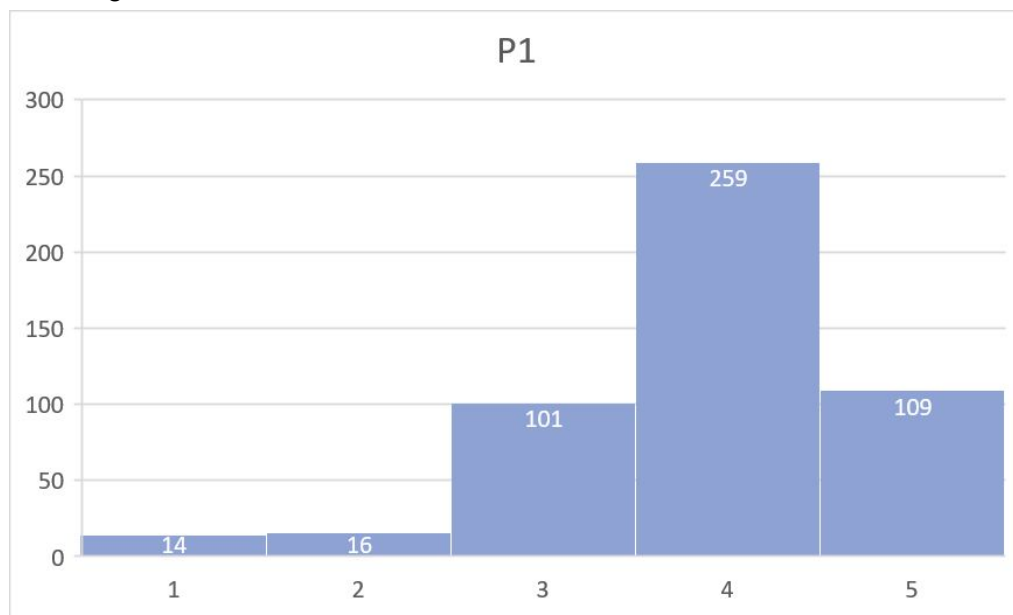
Siguiendo la fórmula del cálculo de puntuación de SUS obtuvimos una media de 70,46 sobre el total de 100 puntos al calcular el promedio de todas las 500 instancias. Esto nos sitúa por encima de la tasa de aceptabilidad de 68 sobre 100 puntos.

Según la teoría sobre los cuestionarios SUS esta puntuación es aceptable, pero nos indica que aún podríamos mejorar nuestro sistema en términos de usabilidad para recibir mejores puntuaciones y aceptación en los usuarios. Si consiguiésemos elevar ese número a 71, se consideraría un buen sistema. Si consiguiésemos elevarlo a 86, sería un sistema excelente.

Ahora pasaremos a analizar los datos de cada pregunta en específico:

**P1** Creo que me gustaría utilizar este sistema frecuentemente:

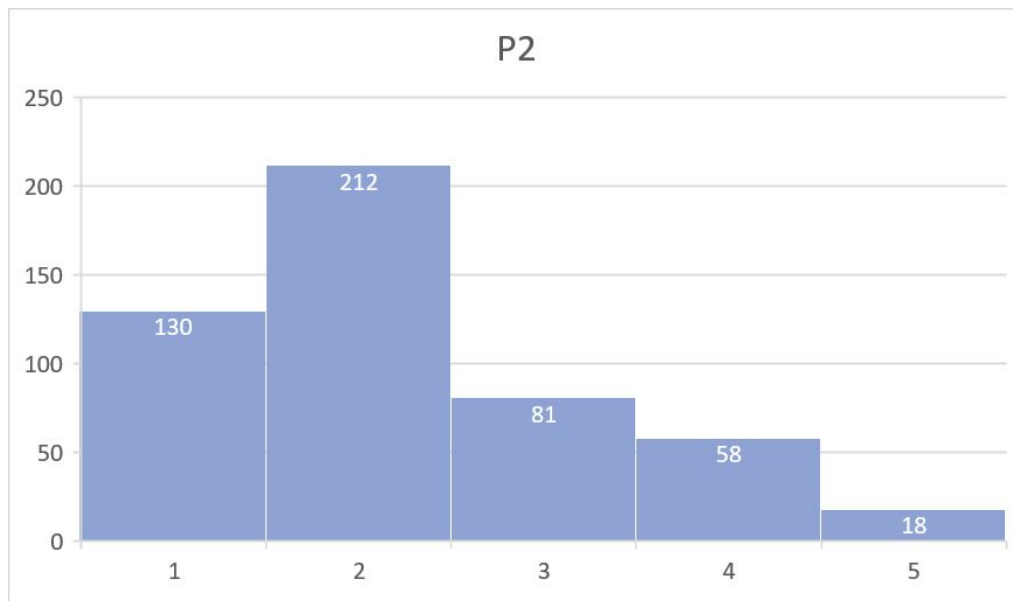
- respuesta promedio: 3,87
- desviación al valor perfecto: 1,13
- gráfico de barras:



- buena distribución cercana al ideal (5)

**P2** Encuentro este sistema innecesariamente complejo:

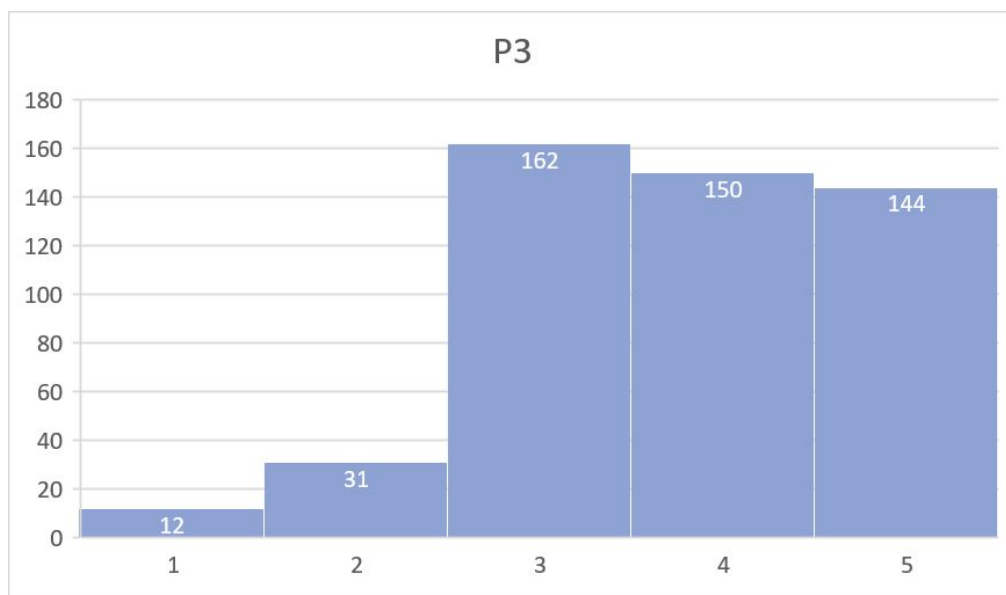
- respuesta promedio: 2,24
- desviación al valor perfecto: 1,24
- gráfico de barras:



- buena distribución cercana al ideal (1)

**P3** Pienso que el sistema es fácil de usar:

- respuesta promedio: 3,77
- desviación al valor perfecto: 1,23
- gráfico de barras:

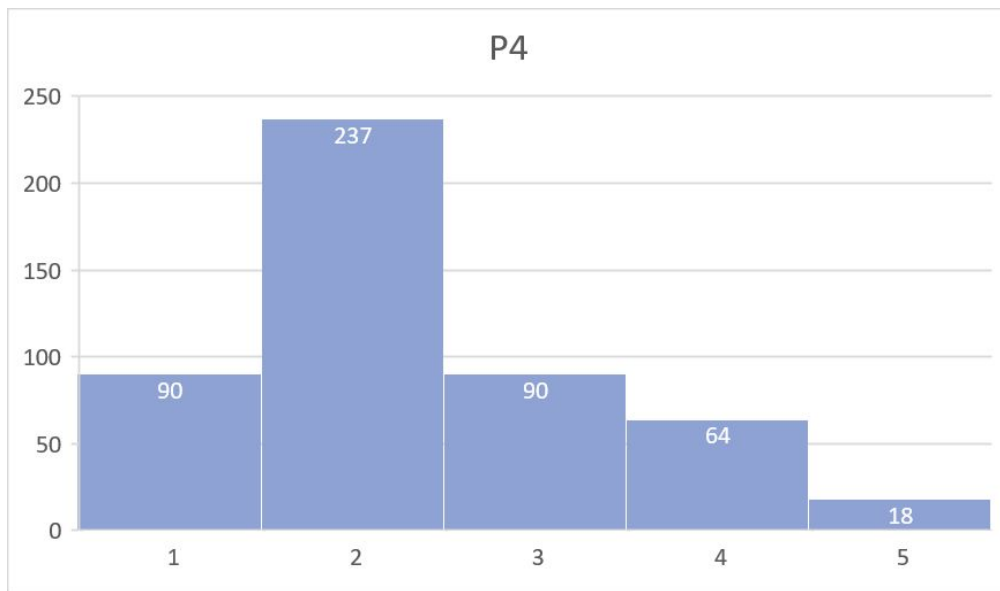


- buena distribución cercana al ideal (5)

**P4** Creo que necesitaría soporte técnico para hacer uso del sistema:

- respuesta promedio: 3,36
- desviación al valor perfecto: 1,36

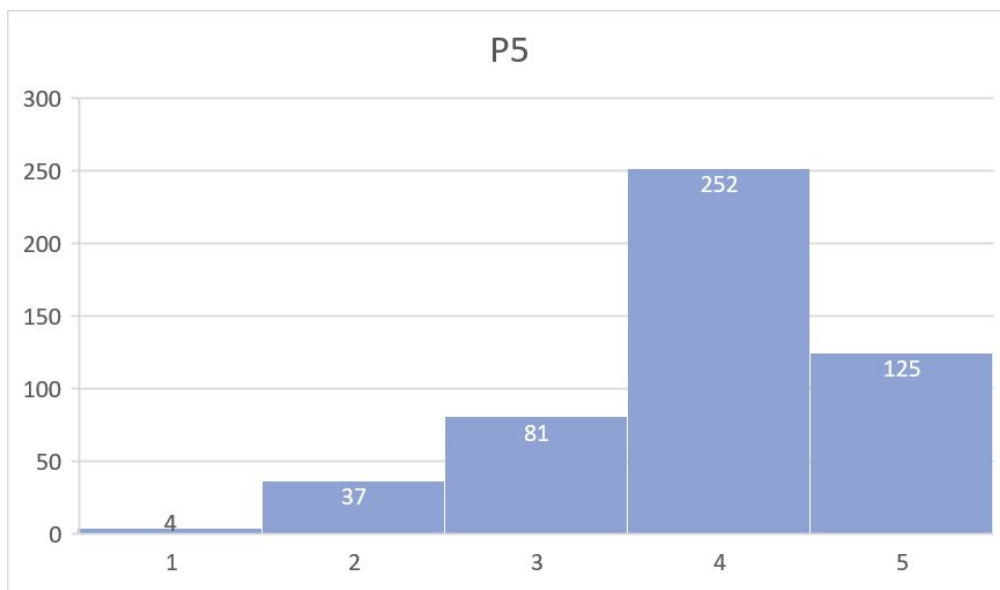
- gráfico de barras:



- buena distribución cercana al ideal (1)

**P5** Encuentro las diversas funciones del sistema bastante bien integradas:

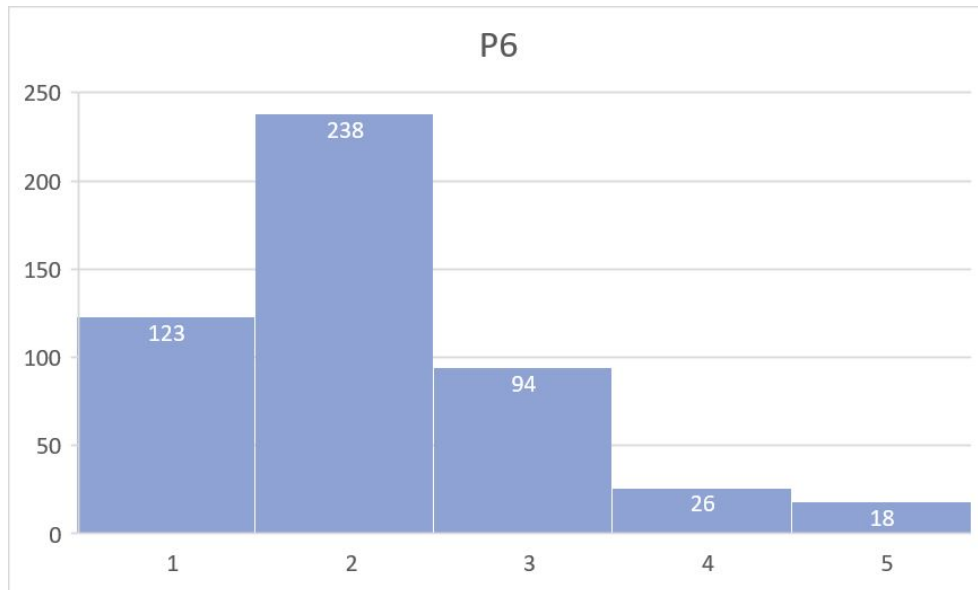
- respuesta promedio: 3,92
- desviación al valor perfecto: 1,08
- gráfico de barras:



- buena distribución cercana al ideal (5)

**P6** He encontrado demasiada inconsistencia en este sistema:

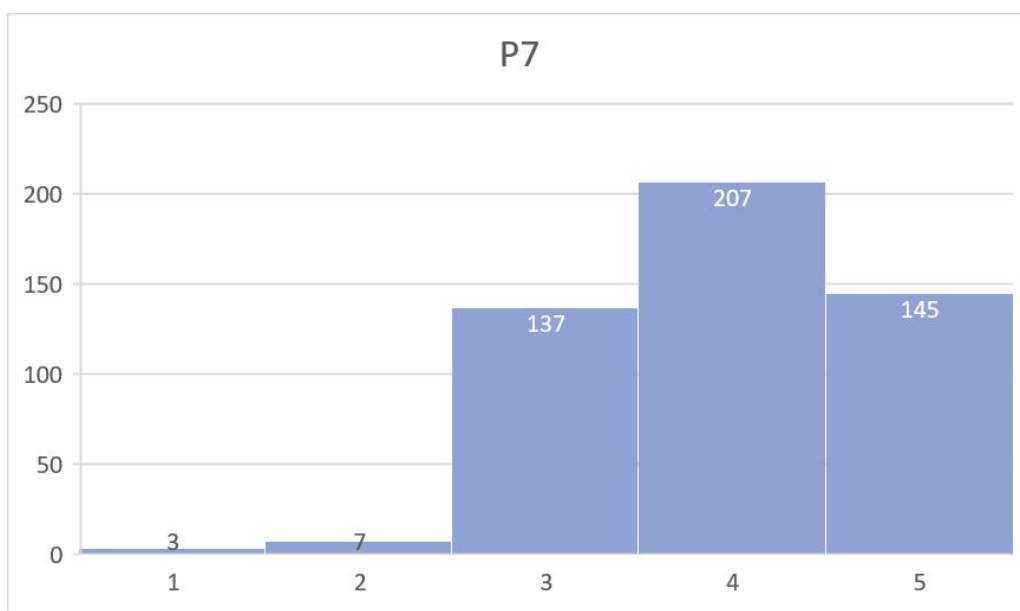
- respuesta promedio: 2,15
- desviación al valor perfecto: 1,15
- gráfico de barras:



- buena distribución cercana al ideal (1)

**P7** Creo que la mayoría de la gente aprendería a hacer uso de este sistema rápidamente:

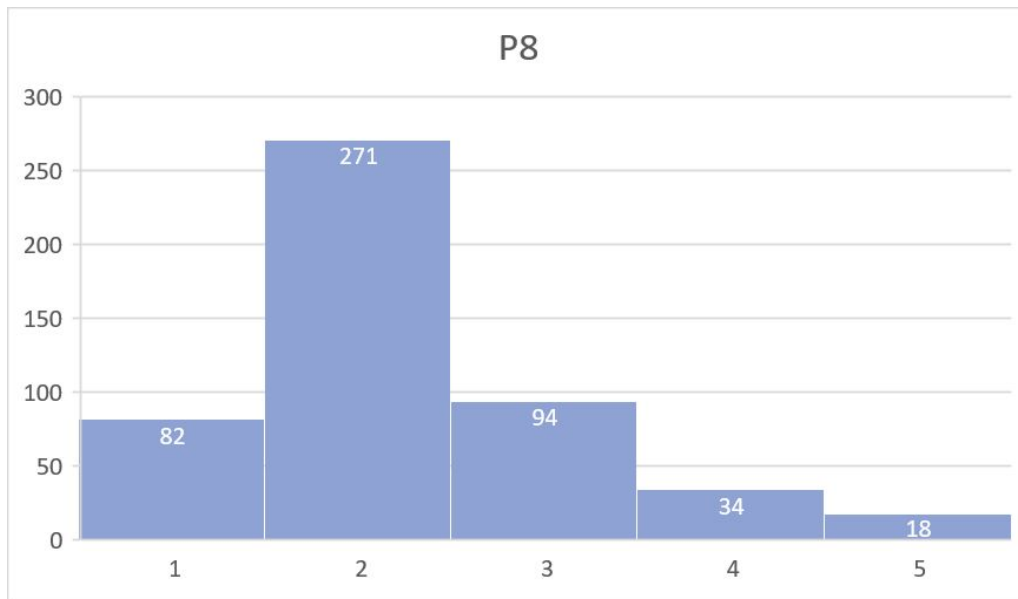
- respuesta promedio: 3,97
- desviación al valor perfecto: 1,03
- gráfico de barras:



- buena distribución cercana al ideal (5)

**P8** He encontrado el sistema bastante incómodo para usar:

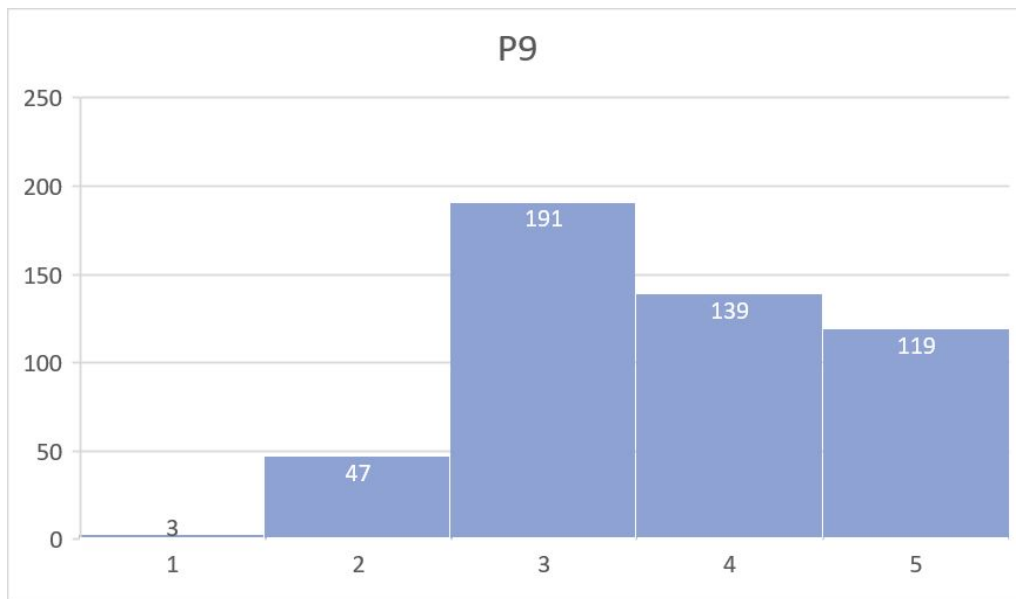
- respuesta promedio: 2,26
- desviación al valor perfecto: 1,26
- gráfico de barras:



- buena distribución cercana al ideal (1)

**P9** Me he sentido muy seguro haciendo uso del sistema:

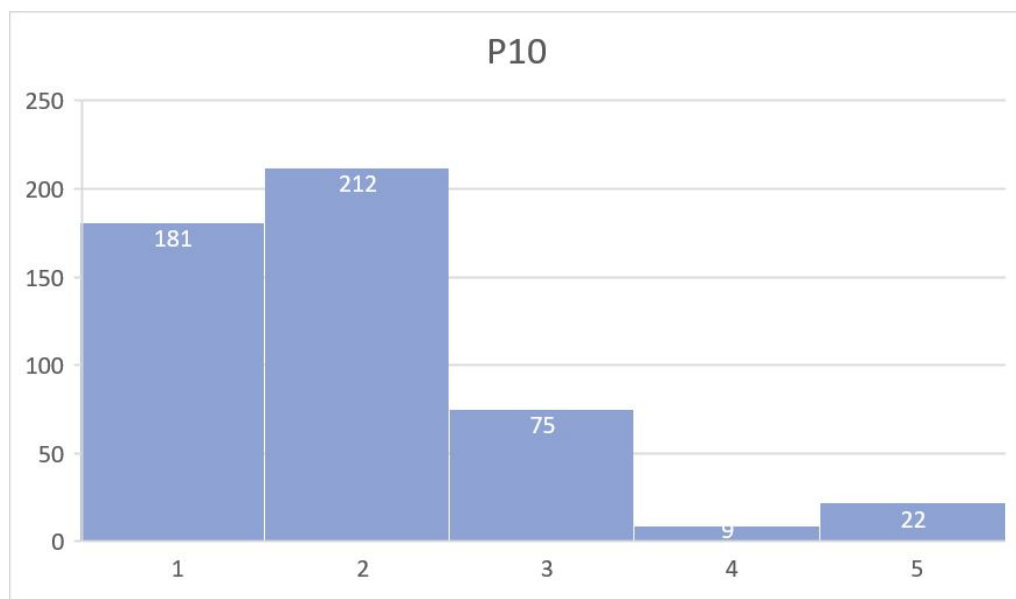
- respuesta promedio: 3,65
- desviación al valor perfecto: 1,35
- gráfico de barras:



- buena distribución cercana al ideal (5)

**P10** Necesitaría aprender un montón de cosas antes de poder manejar el sistema:

- respuesta promedio: 1,95
- desviación al valor perfecto: 0,95
- gráfico de barras:



- muy buena distribución cercana al ideal (1)

### 3. Conclusiones

Como conclusiones al análisis cuantitativo, los resultados son bastante positivos, ya que se tarda relativamente poco en realizar la tarea, y no se necesitan demasiados clicks. Tal vez deberíamos enfocarnos en los errores y los deshacer realizados.

- En cuanto a los errores, estos son bastante consistentes entre los usuarios, por lo que podemos deducir que existen algunas partes de la tarea que son propensas a los mismos errores. Deberíamos investigar cuáles son estas secciones y analizar cómo se podría mejorar el sistema.
- En cuanto al número de deshacer, podemos ver una desviación estándar considerable, que significa que los usuarios tienen distintas maneras de solucionar los errores cometidos. Por lo tanto, sería interesante analizar estas diferencias para facilitar el progreso de la tarea.

Como conclusiones del cuestionario podríamos analizar los datos de cada pregunta para obtener un ránking de las preguntas mejor valoradas según su respuesta ideal gracias a la desviación calculada. Con estos datos, descubrimos que la mejor valorada es la P10 seguida, a continuación, de la P7, después la P5, la P1, la P6, la P3, la P2, la P8, la P9 y por último la P4, que se trata de la peor valorada.

Esto nos indica nuestros puntos fuertes y débiles en cuanto a la usabilidad del sistema. Por ejemplo, nuestros usuarios no consideran que se requiera aprender un montón de cosas antes de poder manejar el sistema (P10), pero algunos consideran que se necesitaría establecer una base de soporte técnico en este (P4).

Basándonos en estas respuestas, podemos analizar cuál sería nuestro foco en la siguiente iteración del diseño, para aplacar primero los problemas que nuestros usuarios consideran más relevantes:

- Añadir alguna forma de soporte técnico, o mejorar la que haya.
- Hacer a los usuarios sentirse más seguros, que podría estar relacionado con la cantidad de vueltas atrás realizadas.
- Hacer al usuario sentirse más cómodo utilizando el sistema.
- Hacer el sistema más sencillo de utilizar.