

Dossier del Proyecto ‘Alcanzando la Cima’ – Análisis Exploratorio de Datos (EDA)
Identificando las Condiciones Óptimas para la Conquista de los Ochomiles del Himalaya



Índice

1. Objetivo Fundamental del Proyecto	3
2. Enfoque y Metodología de Análisis	4
3. Perfiles Detallados de Picos Clave: Cho Oyu y Lhotse	7
3.1. Cho Oyu (8.188m)	9
3.2. Lhotse (8.516m)	10
4. Conclusiones y Recomendaciones Clave	12
5. Gráficas adicionales de consulta	13

1. Objetivo Fundamental del Proyecto

El proyecto 'Alcanzando la Cima' se embarca en un profundo **Análisis Exploratorio de Datos (EDA)** con un objetivo claro y ambicioso: **identificar las condiciones más favorables para poder alcanzar la cima de uno de los 14 picos de más de 8.000 metros del Himalaya**. Este análisis busca ir más allá de la mera estadística para desvelar patrones y correlaciones que puedan informar futuras expediciones.

Para llevar a cabo este estudio exhaustivo, se ha compilado una vasta colección de datos a partir de diversas fuentes, incluyendo exped.csv, himalayan_data_dictionary.csv, members.csv, peaks.csv, y refer.csv. La integración de estos conjuntos de datos ha resultado en un dataset combinado (df) que comprende un total impresionante de **89.153 entradas y 146 columnas**, proporcionando una base de conocimiento robusta para la investigación. La fuente de estos datos es <https://www.kaggle.com/datasets/siddharth0935/himalayan-expeditions/data>.

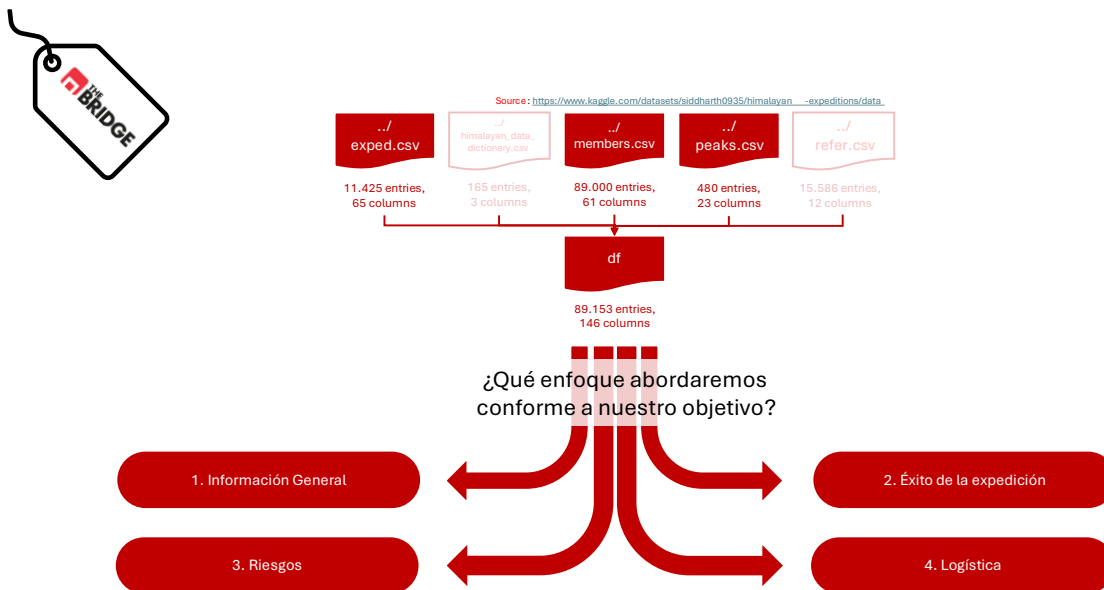


Ilustración 1. Estrategia para analizar la información disponible y responder al objetivo de negocio

2. Enfoque y Metodología de Análisis

Para garantizar que el estudio sea pertinentemente focalizado y ofrezca conclusiones significativas, hemos adoptado un enfoque específico: solo se han considerado aquellos picos que cumplen dos criterios fundamentales:

- Picos con una altura superior a los **8.000 metros**.
- Picos que registran **más de 1.000 expediciones o registros**.

De un total de 14 opciones de picos disponibles para el estudio, este filtro nos permite concentrarnos en aquellos con suficiente volumen de datos para extraer patrones fiables.

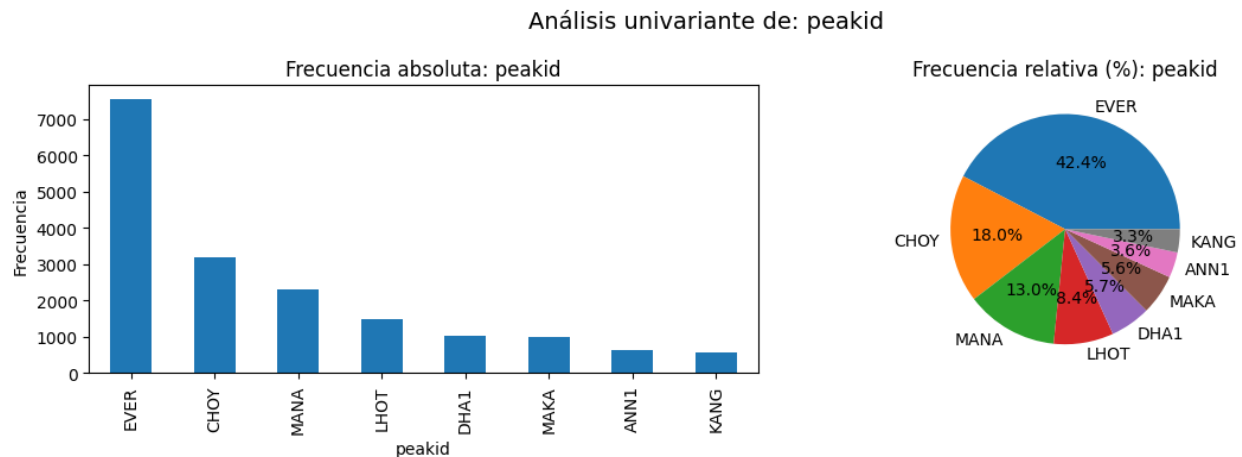


Ilustración 2. Información disponible en el dataset sobre los 8 picos filtrados

El análisis se ha estructurado en cuatro áreas principales, cada una abordando una faceta crítica de las expediciones de alta montaña:

- **1. Información General (df_info):** Este segmento contiene 17.800 entradas y 17 columnas, proporcionando una visión amplia y contextual de las expediciones y los picos.
- **2. Éxito de la Expedición (df_ach):** Con 12.821 entradas y 9 columnas, esta sección es crucial para entender qué factores contribuyen a la consecución del objetivo principal. El concepto de **"éxito absoluto" se define de manera rigurosa como la combinación de llegar a la cima y sobrevivir a la**

experiencia. Se realizan análisis bivariados y multivariados para explorar las interrelaciones entre variables y el éxito.

- **3. Riesgos (df_risk):** Compuesto por 10.425 entradas y 8 columnas, este apartado cuantifica y categoriza los peligros inherentes a estas expediciones. El **"riesgo absoluto" se refiere directamente a la probabilidad de fallecer** durante la expedición. La media registrada de fallecimiento asciende al 4,5%.
- **4. Logística (df_log):** Esta área, con 17.183 entradas y 27 columnas, desglosa los aspectos operativos y de infraestructura que son vitales para la ejecución de una expedición.

Dossier del Proyecto ‘Alcanzando la Cima’ – Análisis Exploratorio de Datos (EDA)
Identificando las Condiciones Óptimas para la Conquista de los Ochomiles del Himalaya
Julio 2.025

<pre>df_inf.info() <class 'pandas.core.frame.DataFrame'> Index: 53542 entries, 52 to 88968 Data columns (total 17 columns): # Column Non-Null Count Dtype --- --- 0 expid_x 53542 non-null object 1 peakid 53542 non-null object 2 host 53542 non-null object 3 nation 53542 non-null object 4 season 53542 non-null object 5 citizen 53542 non-null object 6 myear 53542 non-null int64 7 sherpa 53542 non-null bool 8 tibetan 53542 non-null bool 9 heightm 53542 non-null int64 10 himal 53542 non-null object 11 location 53542 non-null object 12 phost 53542 non-null object 13 region 53542 non-null object 14 pkname 53542 non-null object 15 pseason 53542 non-null object 16 pmonth 53542 non-null object dtypes: bool(2), int64(2), object(13) memory usage: 6.6+ MB</pre>	<pre>df_acg.info() <class 'pandas.core.frame.DataFrame'> Index: 53542 entries, 52 to 88968 Data columns (total 9 columns): # Column Non-Null Count Dtype --- --- 0 expid_x 53542 non-null object 1 peakid 53542 non-null object 2 highpoint 53542 non-null int64 3 death 53542 non-null bool 4 mhighpt 53542 non-null bool 5 mo2none 53542 non-null bool 6 mo2used 53542 non-null bool 7 msuccess 53542 non-null bool 8 psummiters 53542 non-null object dtypes: bool(5), int64(1), object(3) memory usage: 2.3+ MB</pre>
<pre>df_risk.info() <class 'pandas.core.frame.DataFrame'> Index: 53542 entries, 52 to 88968 Data columns (total 8 columns): # Column Non-Null Count Dtype --- --- 0 expid_x 53542 non-null object 1 peakid 53542 non-null object 2 death 53542 non-null bool 3 deathclass 768 non-null object 4 deathdate 750 non-null datetime64[ns] 5 deathhgtm 53542 non-null int64 6 deathtype 768 non-null object 7 hired 53542 non-null bool dtypes: bool(2), datetime64[ns](1), int64(1), object(4) memory usage: 3.0+ MB</pre>	<pre>df_log.info() <class 'pandas.core.frame.DataFrame'> Index: 53542 entries, 52 to 88968 Data columns (total 27 columns): # Column Non-Null Count Dtype --- --- 0 expid_x 53542 non-null object 1 peakid 53542 non-null object 2 agency 47288 non-null object 3 camps 53542 non-null int64 4 campsites 53542 non-null object 5 countries 41644 non-null object 6 leaders 53542 non-null object 7 o2used 53542 non-null bool 8 rope 53542 non-null int64 9 route1 53542 non-null object 10 smtdays 53542 non-null float64 11 sponsor 53542 non-null object 12 success1 53542 non-null bool 13 success2 53542 non-null bool 14 success3 53542 non-null bool 15 success4 53542 non-null bool 16 totmembers 53542 non-null int64 17 tothired 53542 non-null int64 18 bconly 53542 non-null bool 19 mhighpt 53542 non-null bool 20 mroute1 53542 non-null int64 21 mseason 53542 non-null object 22 msuccess 53542 non-null bool 23 sherpa 53542 non-null bool 24 tibetan 53542 non-null bool 25 phost 53542 non-null object 26 region 53542 non-null object dtypes: bool(10), float64(1), int64(5), object(11) memory usage: 7.9+ MB</pre>

Ilustración 3. Información almacenada originalmente en cada uno de los 4 Datasets

3. Perfiles Detallados de Picos Clave: Cho Oyu y Lhotse

El análisis ha permitido perfilar dos de los ochomiles más significativos, el Cho Oyu y el Lhotse, destacando sus particularidades en cuanto a éxito, riesgos y logística. Ambos picos se encuentran en la región de Khumbu-Rowaling-Makalu y están bajo la autoridad de Nepal y China.

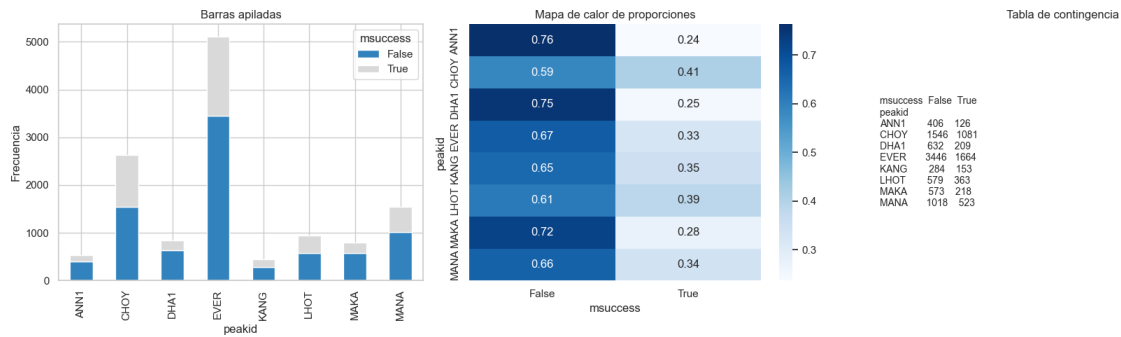


Ilustración 4. Análisis bivariante éxito_expedición(msuccess) vs montañas(peakid)

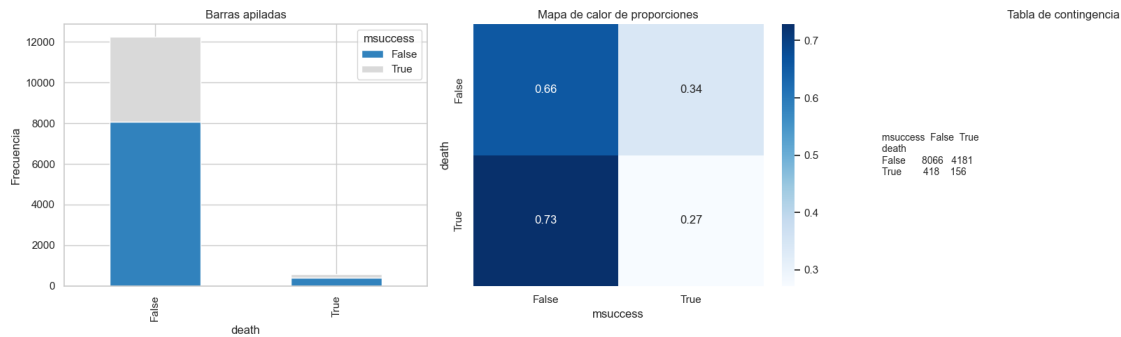


Ilustración 5. Análisis bivariante éxito_expedición(msuccess) vs supervivencia(death)

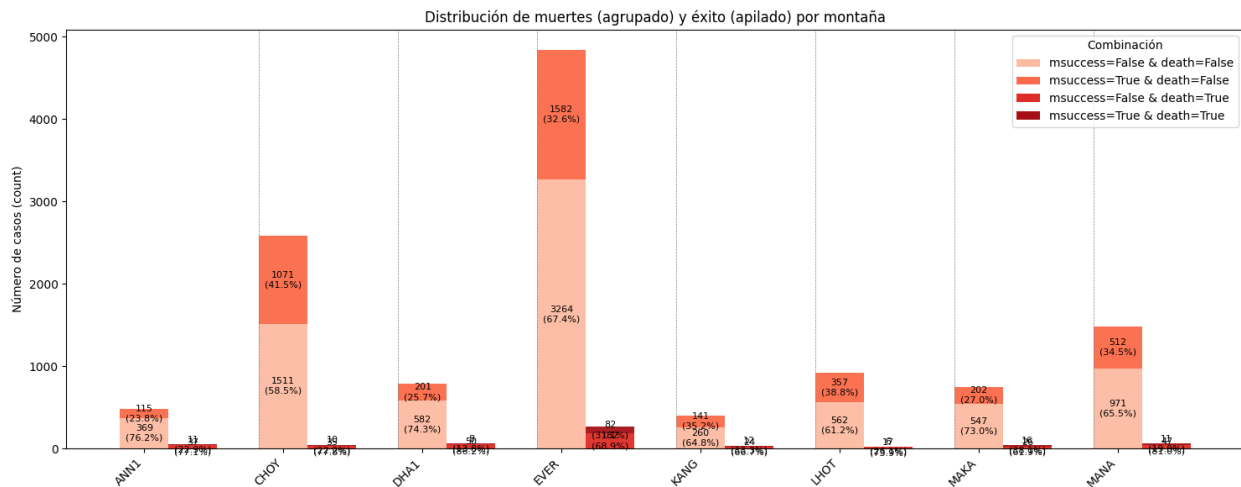


Ilustración 6. Análisis multivariante montañas(peakid), fallecimientos(death) y éxito_expedición(msuccess)

Dossier del Proyecto 'Alcanzando la Cima' – Análisis Exploratorio de Datos (EDA)
Identificando las Condiciones Óptimas para la Conquista de los Ochomiles del Himalaya
Julio 2.025

peakid	death	msuccess	count	% grupo	total_casos
ANN1	False	False	369	76.2	484
ANN1	False	True	115	23.8	484
ANN1	True	False	37	77.1	48
ANN1	True	True	11	22.9	48
CHOY	False	False	1511	58.5	2582
CHOY	False	True	1071	41.5	2582
CHOY	True	False	35	77.8	45
CHOY	True	True	10	22.2	45
DHA1	False	False	582	74.3	783
DHA1	False	True	201	25.7	783
DHA1	True	False	50	86.2	58
DHA1	True	True	8	13.8	58
EVER	False	False	3264	67.4	4846
EVER	False	True	1582	32.6	4846
EVER	True	False	182	68.9	264
EVER	True	True	82	31.1	264
KANG	False	False	260	64.8	401
KANG	False	True	141	35.2	401
KANG	True	False	24	66.7	36
KANG	True	True	12	33.3	36
LHOT	False	False	562	61.2	919
LHOT	False	True	357	38.8	919
LHOT	True	False	17	73.9	23
LHOT	True	True	6	26.1	23
MAKA	False	False	547	73.0	749
MAKA	False	True	202	27.0	749
MAKA	True	False	26	61.9	42
MAKA	True	True	16	38.1	42
MANA	False	False	971	65.5	1483
MANA	False	True	512	34.5	1483
MANA	True	False	47	81.0	58
MANA	True	True	11	19.0	58
peakid	total_fallecidos_peak		% sobre total fallecidos		
EVER	264		46.0		
DHA1	58		10.1		
MANA	58		10.1		
ANN1	48		8.4		
CHOY	45		7.8		
MAKA	42		7.3		
KANG	36		6.3		
LHOT	23		4.0		

Ilustración 7. Datos tabulados análisis multivariante

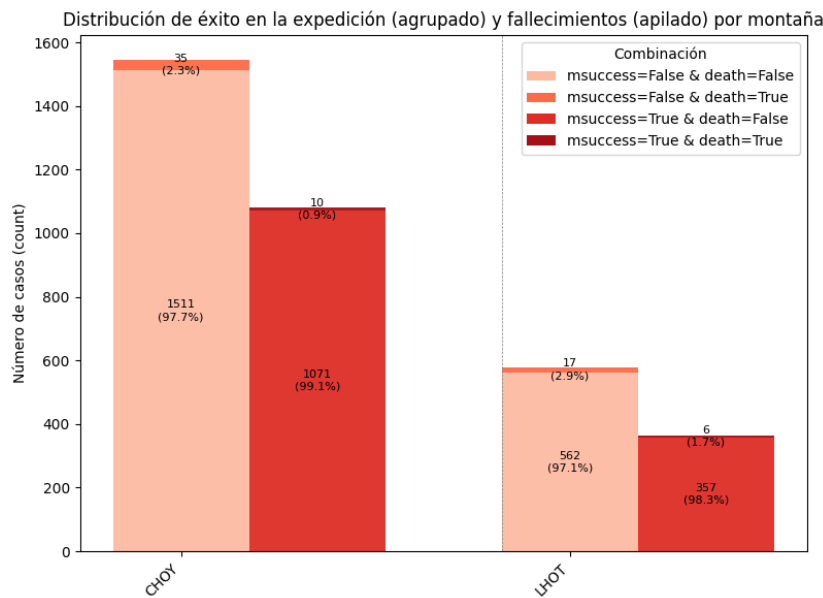


Ilustración 8. Información mutivariante limitada a las dos montañas de interés, Cho Oyu (CHOY) y Lhotse (LHOT)

3.1. Cho Oyu (8.188m)

Considerado a menudo como uno de los ochomiles "más accesibles", el Cho Oyu presenta unas estadísticas de éxito y riesgo que lo hacen particularmente interesante:

- **Tasa de Éxito:** Un notable **41%** de las expediciones alcanzan la cima y sus miembros regresan con vida.
- **Tasa de Fallecimiento:** La probabilidad de fallecer en el Cho Oyu es del **2%**, significativamente inferior a la media general de los ochomiles.
- **Causas de Muerte Más Frecuentes:** Las principales causas de mortalidad son el **descenso desde el intento de cima (36,17%)** y la **preparación de ruta (36,17%)**. Otras causas incluyen la muerte en el Campo Base (BC) o Campo Base Avanzado (ABC) (10,64%), la evacuación de la expedición (10,64%), el ascenso en el intento de cima (4,26%) y la muerte en ruta hacia el BC (2,13%).
- **Logística Clave para el Éxito en Cho Oyu:**
 - Existe **1 agencia especializada** en expediciones a este pico.
 - La infraestructura incluye un Campo Base (BC) y **cuatro campamentos adicionales** a diferentes altitudes.
 - Se ha observado un **mayor porcentaje de éxito cuando se utiliza cuerda y oxígeno suplementario**.

- La expedición típica se realiza por el **vértice Oeste** y está programada para una duración de **22 días**.
- No es posible acampar solo en el Campo Base.
- Dispone de **dos rutas alternativas** para el ascenso.
- Las tasas de éxito son similares en Primavera (30%) y Verano (33%), pero el **Otoño emerge como la estación con el mayor porcentaje de éxito (43%)**.
- Una recomendación logística crucial es **contratar un sherpa tibetano, ya que las expediciones con su apoyo tienen un 64% de éxito**.

3.2. Lhotse (8.516m)

Vecino del Everest, el Lhotse es un desafío formidable por derecho propio, ofreciendo una perspectiva distinta en términos de éxito y logística:

- **Tasa de Éxito:** Un **38%** de las expediciones al Lhotse culminan con éxito y supervivencia.
- **Tasa de Fallecimiento:** Similar al Cho Oyu, la tasa de fallecimiento registrada es del **2%**.
- **Causas de Muerte Más Frecuentes:** La causa predominante de fatalidad es el **descenso desde el intento de cima (52,17%)**. Otras causas notables son la preparación de ruta (26,09%), el ascenso en el intento de cima (17,39%) y la muerte en el BC/ABC (4,35%).
- **Logística Clave para el Éxito en Lhotse:**
 - A diferencia del Cho Oyu, el Lhotse **no cuenta con una agencia especializada**.
 - También dispone de un Campo Base (BC) y **cuatro campamentos adicionales**.
 - El éxito se asocia con el uso de **oxígeno y la ausencia de cuerda** en ciertas secciones.
 - La expedición se lleva a cabo principalmente por la **cara Oeste** y suele tener una duración programada de **45 días**.
 - No es posible acampar solo en el Campo Base.
 - Se ha identificado que la **ruta 2 aumenta el porcentaje de éxito**.

- Los porcentajes de éxito **aumentan significativamente en Primavera**.
- Es imprescindible **contratar sherpas** para esta expedición.

4. Conclusiones y Recomendaciones Clave

El análisis detallado de picos como Cho Oyu y Lhotse subraya que el éxito en una expedición a un ochomil va mucho más allá de la preparación física del alpinista. Es una compleja interacción de factores logísticos y estratégicos.

Para maximizar las posibilidades de alcanzar la cima y, crucialmente, sobrevivir a la experiencia, se deben considerar los siguientes elementos:

- **Elección de la Ruta y Estación:** Cada pico tiene sus "ventanas" óptimas y rutas preferidas que pueden influir drásticamente en el éxito.
- **Uso de Equipo Esencial:** La decisión sobre el uso de cuerda y oxígeno debe basarse en el pico específico y sus condiciones.
- **Apoyo Profesional:** La contratación de sherpas experimentados es un factor determinante, especialmente en picos como el Cho Oyu (con un 64% de éxito con sherpa tibetano) y el Lhotse, donde su presencia es fundamental.
- **Conocimiento de Riesgos:** Comprender las causas más frecuentes de fatalidad (ej. descenso desde la cima, preparación de ruta) permite una mejor planificación y toma de decisiones para mitigar estos riesgos.

Piense en la planificación de una expedición a un pico de 8.000 metros como la **construcción de un edificio imponente**: cada fase, desde la cimentación (información general y análisis de datos) y la estructura (planificación logística), hasta la seguridad de los trabajadores (mitigación de riesgos y apoyo de sherpas), debe ser meticulosamente diseñada y ejecutada. El objetivo final, la cima, es el coronamiento de un esfuerzo donde cada detalle, por pequeño que parezca, contribuye a la solidez y el éxito de la obra completa.

Lamentablemente el dataset permitía análisis bi y multivariante entre variables categóricas limitando nuestra capacidad de análisis.

5. Gráficas adicionales de consulta

