

1.

Do you think AI will replace human financial analysts in the near future?

AI is unlikely to completely replace human financial analysts in the near future, but it will significantly reshape their roles. AI excels at processing large datasets, detecting patterns, and automating routine tasks like financial modeling and risk assessment. However, human analysts are still essential for interpreting complex economic trends, making strategic decisions, and considering qualitative factors like market sentiment and geopolitical risks.

Moreover, AI models lack transparency ("black box" problem) and can struggle with unpredictable financial crises that deviate from historical data. Regulatory and ethical concerns also require human oversight. Instead of full replacement, the future of finance will likely involve AI-human collaboration, where AI enhances efficiency while analysts focus on high-level decision-making and nuanced judgment.

How is AI currently used in the finance industry, particularly in risk assessment and fraud detection?

AI plays a crucial role in finance, especially in **risk assessment** and **fraud detection**:

1. Risk Assessment

- **Credit Scoring & Loan Decisions:** AI models analyze traditional financial data (e.g., credit history, income) alongside alternative data (e.g., spending patterns, online activity) to assess creditworthiness.
- **Market Risk Analysis:** AI-powered predictive models help banks and investment firms assess potential risks in stocks, bonds, and derivatives by identifying patterns in historical data.
- **Algorithmic Trading:** AI algorithms analyze vast datasets to optimize trading strategies and minimize risk exposure.

What are some limitations of AI in financial decision-making?

AI has several limitations in financial decision-making, including:


1. Data Quality and Bias

- AI models rely on large datasets, and if the data is incomplete, biased, or outdated, predictions can be inaccurate.
- Historical financial data may not fully capture emerging risks or market anomalies.

2. Lack of Explainability ("Black Box" Problem)

- Many advanced AI models, especially deep learning algorithms, are difficult to interpret.
- This makes it challenging to justify decisions to regulators, clients, or internal auditors.

3. Difficulty Handling Unpredictable Events

- AI struggles with rare, high-impact financial crises (e.g., the 2008 financial crash or COVID-19 market disruptions) because these events lack sufficient historical data for training.
- Human intuition and experience are still  ntial for navigating unprecedented situations.

I pose these questions to understand the benefits and challenges of AI in finance. This is a question related to career planning, as I want to pursue a finance-related career afterward, and I hope that ChatGPT's answers will help me.

The responses were well organized and informative, correctly pointing out AI's strengths in data processing and risk analysis, while also acknowledging its limitations in terms of interpretability and dealing with uncertainty. However, I might question the extent to which AI will remain dependent on human oversight as models improve. This raises thought-provoking questions about the evolving relationship between AI and financial professionals.

2.

Creating $k=2$ clusters by applying Lloyd's algorithm.

$$C_1 = \{1, 4\} \quad C_2 = \{2, 3, 5\}$$

$$\mu_j^{C_j} = \frac{1}{\#C_j} \sum_{i \in C_j} x_{ij} \text{ for each } j=1, \dots, p.$$

μ^{C_k} :

$$\mu^{C_1} = (\mu_1^{C_1}, \mu_2^{C_1}) \quad \mu_1^{C_1} = \frac{1}{2}(0+3) = \frac{3}{2} \quad \mu_2^{C_1} = \frac{1}{2}(3+1) = 2$$

$$= (\frac{3}{2}, 2)$$

$$\mu^{C_2} = (\mu_1^{C_2}, \mu_2^{C_2}) \quad \mu_1^{C_2} = \frac{1}{3}(1+1+4) = 2 \quad \mu_2^{C_2} = \frac{1}{3}(3+2+2) = \frac{7}{3}$$

$$= (2, \frac{7}{3})$$

row	column 1	column 2	$C_1(\mu)$	$C_2(\mu)$
1	0	3	$\sqrt{(0-\frac{3}{2})^2 + (3-2)^2} = 1.8$	$\sqrt{(0-2)^2 + (3-\frac{7}{3})^2} = 2.1$
2	1	3	$\sqrt{(1-\frac{3}{2})^2 + (3-2)^2} = 1.1$	$\sqrt{(1-2)^2 + (3-\frac{7}{3})^2} = 1.2$
3	1	2	$\sqrt{(1-\frac{3}{2})^2 + (2-2)^2} = 0.5$	$\sqrt{(1-2)^2 + (2-\frac{7}{3})^2} = 1.1$
4	3	1	$\sqrt{(3-\frac{3}{2})^2 + (1-2)^2} = 1.8$	$\sqrt{(3-2)^2 + (1-\frac{7}{3})^2} = 1.7$
5	4	2	$\sqrt{(4-\frac{3}{2})^2 + (2-2)^2} = 2.5$	$\sqrt{(4-2)^2 + (2-\frac{7}{3})^2} = 2$

$$\Rightarrow \hat{C}_1 = \{1, 2, 3\} \neq C_1 \quad \hat{C}_2 = \{4, 5\} \neq C_2$$

$$\mu^{C_1} = (\frac{1}{3}(0+1+1), \frac{1}{3}(3+3+2)) = (\frac{2}{3}, \frac{8}{3})$$

$$\mu^{C_2} = (\frac{1}{2}(3+4), \frac{1}{2}(1+2)) = (3, \frac{3}{2})$$

$\sqrt{(x_1 - \mu_1)^2 + (x_2 - \mu_2)^2} = \text{distance}$

row	column 1	column 2	$C_1(\mu)$	$C_2(\mu)$
1	0	3	0.7	3.8
2	1	3	0.5	2.9
3	1	2	0.7	2.5
4	3	1	2.9	0.7
5	4	2	3.4	0.7

$$\Rightarrow \bar{C}_1 = \{1, 2, 3\} = \hat{C}_1 \quad \bar{C}_2 = \{4, 5\} = \hat{C}_2$$

Lloyd algorithm terminates, optimal pair of clusters is $C_1 = \{1, 2, 3\}$
 $C_2 = \{4, 5\}$

3.

```
data <- as.matrix(read.table("~/Downloads/Mark_data.txt", header = FALSE))
```

```
data
```

```
Kmax <- 10
```

```
ss <- rep(0, Kmax)
```

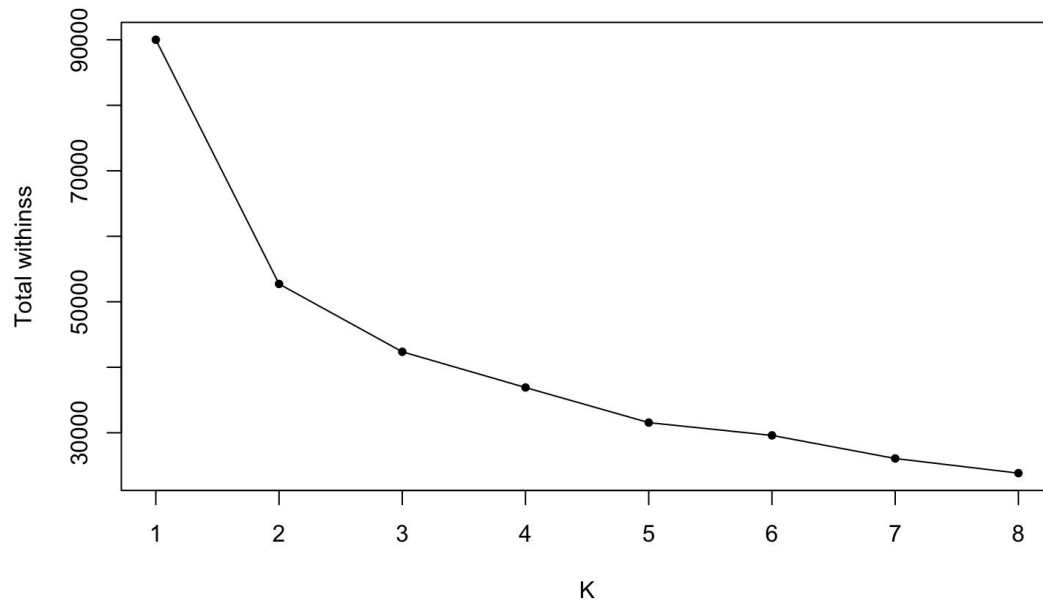
```
for (K in 1:Kmax) {
```

```
  kmeansResult <- kmeans(data, K, iter.max = 100, nstart = 20, algorithm = "Lloyd")
```

```

ss[K] <- kmeansResult$tot.withinss
}
plot(ss, type="o", xlab="K", ylab="Total withinss", pch=20)

```



The 'elbow method' tells us to look for a value of K where steep decrease changes to slow decrease. Looking at the plot above we can say that $K = 3$. Hence, the optimal number of clusters is 3.

```

4.
install.packages("copula")
library(copula)
n<-5000
par(mfrow=c(2,2))
cop<-frankCopula(-700)
distr<-mvdc(cop,margins=c("gamma","gamma"),
            paramMargins=list(list(2,2),
                                list(2,2)))

samples<-rMvdc(n,distr)
print(plot(samples,xlab="x",ylab="y"))
cop<-frankCopula(700)
distr<-mvdc(cop,margins=c("gamma","gamma"),
            paramMargins=list(list(2,2),
                                list(2,2)))

samples<-rMvdc(n,distr)
print(plot(samples,xlab="x",ylab="y"))

```

```

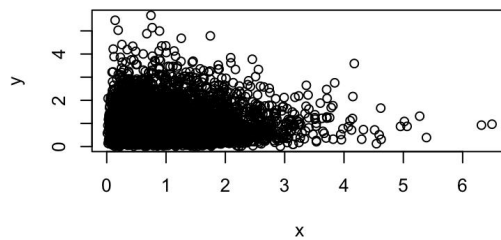
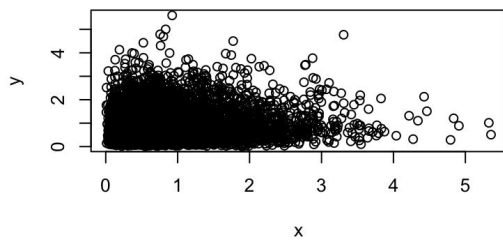
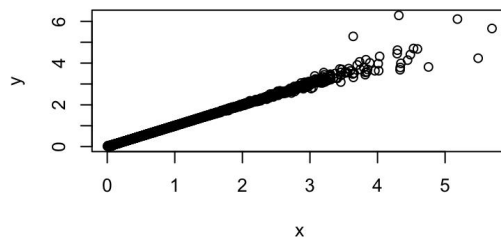
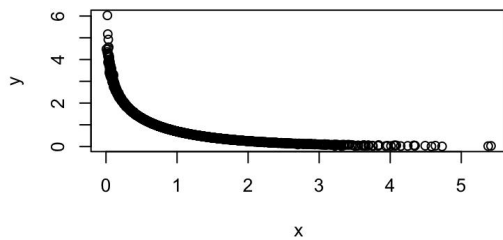
cop<-frankCopula(-0.01)
distr<-mvdc(cop,margins=c("gamma","gamma"),
            paramMargins=list(list(2,2),
                                list(2,2)))

samples<-rMvdc(n,distr)
print(plot(samples,xlab="x",ylab="y"))

cop<-frankCopula(0.01)
distr<-mvdc(cop,margins=c("gamma","gamma"),
            paramMargins=list(list(2,2),
                                list(2,2)))

samples<-rMvdc(n,distr)
print(plot(samples,xlab="x",ylab="y"))

```



4

(

b

)

Q4 (b)

$$\theta \rightarrow -\infty, e^{-\theta} = e^{|\theta|} \Rightarrow \lim_{\theta \rightarrow -\infty} C_{\theta}(u, v) = -\frac{1}{\theta} \log \left(1 + \frac{(e^{|\theta|u}-1)(e^{|\theta|v}-1)}{e^{|\theta|}-1} \right)$$

$$\approx -\frac{1}{\theta} \log \left(1 + \frac{e^{|\theta|(u+v)}}{e^{|\theta|}} \right) = -\frac{1}{\theta} \log \left(1 + e^{|\theta|(u+v-1)} \right)$$

$$\text{if } |\theta|(u+v-1) < 0, -\frac{1}{\theta} \log \left(1 + e^{|\theta|(u+v-1)} \right) \approx -\frac{1}{\theta} e^{|\theta|(u+v-1)} \rightarrow 0$$

$$\text{if } |\theta|(u+v-1) > 0, -\frac{1}{\theta} \log \left(1 + e^{|\theta|(u+v-1)} \right) \approx -\frac{1}{\theta} |\theta|(u+v-1) \rightarrow u+v-1$$

As $\theta \rightarrow -\infty$, the Frank copula converges to:

$$C_{\theta}(u, v) \rightarrow \max(u+v-1, 0)$$

This is the countermonotonicity copula, which represents perfect negative dependence. It means that if one variable is high, the other must be low, like in the case of perfectly negatively correlated variables. This can be seen from the plot in (a) where $\theta = -700$ where there is a downward sloping pattern.

$\theta \rightarrow 0, e^{-\theta x} \approx 1 - \theta x$ using Taylor's approximation

$$\lim_{\theta \rightarrow 0} C_{\theta}(u, v) = -\frac{1}{\theta} \log \left(1 + \frac{(1-\theta u-1)(1-\theta v-1)}{1-\theta-1} \right)$$

$$= -\frac{1}{\theta} \log(1 - \theta uv) \quad \text{since } \log(1+x) \approx x \approx -\frac{1}{\theta}(\theta uv) = uv$$

The independence copula means that variables X and Y become independent in the limit. There is no correlation or dependence between the two random variables. Thus, the limiting case of the Frank copula for $\theta \rightarrow 0$ is simply the product copula: $C(u, v) = uv$. This can be seen from the plot in (a) where $\theta = 0.01$ or -0.01 where there is a downward sloping pattern.

$$\theta \rightarrow \infty, e^{-\theta x} \rightarrow 0, \lim_{\theta \rightarrow \infty} C_{\theta}(u, v) = -\frac{1}{\theta} \log \left(1 + \frac{(0-1)(0-1)}{0-1} \right) = -\frac{1}{\theta} \log(0) \rightarrow \min(u, v)$$

As $\theta \rightarrow \infty$, the Frank copula converges to the comonotonicity ~~upper~~ copula:

$$C(u, v) = \min(u, v)$$

This corresponds to perfect positive dependence, meaning that the two variables are completely dependent. This we can see from plot in (a) where $\theta = 700$ where there is a downward sloping pattern.