	CS 434 HWI	
	TO TO THE TIME	Gretel Rajamoney
QI:	the likelihood function of a given	a dataset D= {x, x,, x, } is
	$P(D \lambda) = \prod_{i=1}^{N} e^{-\lambda} \frac{\lambda^{x_i}}{(x_i)!} = e^{-N\lambda} \frac{\lambda}{N}$	Σ χ; (ε)
	(xi)! - A	(x_i, j)
	the log-likelihood function of \ is fo	und by log (P(DIX)) and is:
	109eP(DIX) = 109e (NX) = 1/1 (x;!)	
	= 10ge(e-NX) + 10ge(x=x)-100	(H(v.1))
	= -N2 + \\[\int \xi \left(\log_e \lambda \right) - \\ \int \log_e \left(\xi \!)	Je (E)(VI)
	next we take the derivative with re	espect to parameter \(\lambda\):
	d [10ge P(DIX)] = d [-NX +]	
		The constant of the control of the c
	MAXIS OF ELL HAT THE WAR WAR AND THE REAL PROPERTY OF THE PARTY OF THE	
0	next we solve for the maximizing vo $-N + \frac{1}{2} \sum_{i=0}^{N} x_i = 0$	Hue $\hat{\lambda}_{ME}$, $\frac{d}{d\lambda} [\log_e(P(D \lambda))] = 0$:
	$=\frac{1}{3}\left(-N\lambda+\sum_{i=1}^{N}x_{i}\right)=0$	
	$\lambda = \frac{1}{N} \sum_{i=1}^{N} x_i$	and sowe for 3
	$\lambda = X$	ALL
_	therefore, $\hat{\lambda}_{MLE} = \overline{X}$	
	since in poisson distribution & is	
	occurrences in an interval, we con	
	2 can be estimated by the mean	number of occurrences
	within the sample data	200
	SHE MAKIMUM O DOSHOLOTU BERMARKE	CON A IS COMMENT
0		

Q2: the log posterior of λ is $\log(P(\lambda D)) \propto \log(P(D \lambda)) + \log(P(\lambda))$ $f(x) = \frac{e^{\lambda \beta}}{e^{\lambda}} \cdot \frac{\lambda^{\alpha}}{e^{\alpha}} \cdot \frac{e^{\alpha}}{e^{\alpha}} \cdot \frac{e^{\alpha}}{e^{\alpha}} \cdot \frac{e^{\alpha}}{e^{\alpha}} \cdot \frac{\lambda^{\alpha}}{e^{\alpha}} \cdot \frac{e^{\alpha}}{e^{\alpha}} \cdot e^{\alpha$		
the log posterior of λ is $\log(P(\lambda D)) \propto \log(P(D \lambda)) + \log(P(\lambda))$ $f(x) = e^{\lambda \frac{1}{2}} \cdot \lambda^{\frac{1}{2}} \cdot \frac{1}{2} \times \lambda^{\frac{1}{2}} \cdot \frac{1}{2} \times \lambda^{\frac{1}{2}} \cdot \frac{1}{2} \times $		
the log posterior of λ is $\log(P(\lambda D)) \propto \log(P(D \lambda)) + \log(P(\lambda))$ $f(x) = e^{\lambda \frac{1}{2}} \cdot \lambda^{\frac{1}{2}} \cdot \frac{1}{2} \times \lambda^{\frac{1}{2}} \cdot \frac{1}{2} \times \lambda^{\frac{1}{2}} \cdot \frac{1}{2} \times $		
$f(x_{i}, \lambda) = \prod_{i=1}^{N} f(x_{i} \lambda) f(x)$ $= e^{-N\lambda} \lambda_{i}^{2} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= e^{-N\lambda} \lambda_{i}^{2} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= e^{-N\lambda} \lambda_{i}^{2} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{3^{\alpha}}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{3^{\alpha}}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{-\lambda \beta} \lambda^{\alpha-1} e^{-\lambda \beta} \lambda^{\alpha-1} e^{-\lambda \beta} e^{-\lambda \beta} \lambda^{\alpha-1} $		
$f(x_{i}, \lambda) = \prod_{i=1}^{N} f(x_{i} \lambda) f(x)$ $= e^{-N\lambda} \lambda_{i}^{2} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= e^{-N\lambda} \lambda_{i}^{2} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= e^{-N\lambda} \lambda_{i}^{2} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{3^{\alpha}}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{3^{\alpha}}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{-\lambda \beta} \lambda^{\alpha-1} e^{\alpha}$ $= \int_{0}^{\infty} \frac{1!}{1!} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} e^{-\lambda \beta} \lambda^{\alpha-1} e^{-\lambda \beta} \lambda^{\alpha-1} e^{-\lambda \beta} e^{-\lambda \beta} \lambda^{\alpha-1} $	Q2:	the log posterior of 2 is inclosed as a local
$f(x_{i}, \lambda) = \inf_{x \in \mathbb{N}} f(x_{i} \lambda) f(x)$ $= e^{-N\lambda} \lambda_{i}^{2} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} \beta^{\alpha}$ $= e^{-N\lambda} \lambda_{i}^{2} x_{i} = e^{-\lambda \beta} \lambda^{\alpha-1} \beta^{\alpha}$ $= e^{-N\lambda} \lambda_{i}^{2} x_{i} + \alpha^{-1} d$ $= f(x) \lambda_{i}^{2} \lambda_{i}$ $= f(x) \lambda_{i}^{2} \lambda_{i}$ $= f(x) \lambda_{i}^{2} \lambda_{i}^{2} \lambda_{i} + \alpha^{-1} d$ $= f(x) \lambda_{i}^{2} \lambda_{i}^{2} \lambda_{i} + \alpha^{-1} d$ $= e^{-N\lambda} \lambda_{i}^{2} \lambda_{i}^$		
The maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$ The maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$		(0)
The maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$ The maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$		$f(x_i, \lambda) = \prod_{i=1}^{n} f(x_i \lambda) f(x)$
The maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$ The maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$		$= e^{-N\lambda} \lambda^{\frac{N}{N}} x_1 - e^{-\lambda \beta} \lambda^{\alpha-1} e^{-\lambda}$
The maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$ The maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$		$\prod_{i=1}^{n} \lfloor x_i \rfloor \qquad P(\infty)$
The maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$ The maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$	- 11	Ba -NX - XB ZXi 20-1 d
The maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$ The maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$		Jo Tikila
The maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$ The maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$		$\frac{1}{2}$ $\frac{1}$
The maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$ The maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$		(x) 2 (x) 2 (x) 3 (x)
$\frac{d}{d\lambda} \cdot P(\lambda \mid D) = 0 + \frac{d}{d\lambda} \left[\lambda(N + B) \right] + \frac{d}{d\lambda} \left\{ \sum_{i=1}^{N} x_i + \alpha - 1 \right\}$ $= -(N + B) + \sum_{i=1}^{N} x_i + \alpha - 1$ $= -(N + B) + \sum_{i=1}^{N} x_i + \alpha - 1$ $= \sum_{i=1}^{N} x_i + \alpha - 1 = \lambda(N + B)$ $= \sum_{i=1}^{N} x_i + \alpha - 1$ $= N + B$ $\lambda = \lambda$ the maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$		$p(x 0) = \frac{1}{f(x)}$
now we set the derivative to zero and solve for $\hat{\lambda}_{MAP}$ $= (N+B) + \frac{1}{12} \frac{x_i + \alpha - 1}{\lambda} = 0$ $= \frac{1}{12} x_i + \alpha - 1 = \lambda(N+B)$ $= \frac{1}{12} x_i + \alpha - 1$ $= N+B$ $\lambda = \hat{\lambda}$ the maximum a posteriori estimator of $\hat{\lambda}$ is $\frac{\sum_{i=1}^{N} x_i + \alpha - 1}{\sum_{i=1}^{N} x_i + \alpha - 1}$		
now we set the derivative to zero and solve for $\hat{\lambda}_{MAP}$ $= \sum_{i=1}^{N} x_i + \alpha - 1 = \lambda(N+B)$ $= \sum_{i=1}^{N} x_i + \alpha - 1 = \lambda(N+B)$ $= \lambda + \lambda$ the maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$		ax P(XID) = 0+ ax [x(N+B)]+ ax (h=x) [xx+a-1]
now we set the derivative to zero and solve for $\hat{\lambda}_{MAP}$ $= \sum_{i=1}^{N} x_i + \alpha - 1 = \lambda(N+B)$ $= \sum_{i=1}^{N} x_i + \alpha - 1 = \lambda(N+B)$ $= \lambda + \lambda$ the maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$		Z X i + oc - 1
$= \sum_{i=1}^{N} x_i + \alpha - 1 = \lambda(N+B)$ $= \sum_{i=1}^{N} x_i + \alpha - 1$ the maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$		a (Myth) water
$= \sum_{i=1}^{N} x_i + \alpha - 1 = \lambda(N+B)$ $= \sum_{i=1}^{N} x_i + \alpha - 1$ the maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$		now we set the derivative to zero and solve for 2
$= \sum_{i=1}^{N} x_i + \alpha - 1 = \lambda(N+B)$ $= \sum_{i=1}^{N} x_i + \alpha - 1$ $= \frac{1}{N+B}$ $\lambda = \lambda$ the maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$		- (N+8) + 12 x; + a-1
$\lambda = \lambda$ the maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$		^
$\lambda = \hat{\lambda}$ the maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$		N.
$\lambda = \hat{\lambda}$ the maximum a posteriori estimator of λ is $\sum_{i=1}^{N} x_i + \alpha - 1$		181
the maximum a posteriori estimator of λ is $\sum_{i=1}^{\infty} x_i + \alpha - 1$		
the maximum a posteriori estimator of λ is $\frac{1}{N+8}$		y = y
		the maximum a posteriori estimator of \(\lambda\) is \(\frac{1}{2}\)\(\frac{1}{2}\)

	S. Levinster 21 - Linear Algenta & Linear Regression)
Q3:	the conditional interest to
ì	the conditional joint density of y, y,, y, is:
	f(y,, y2,, yn/x) = \frac{1}{11} \left(\frac{e^{-\lambda}}{2}, \lambda \frac{\pi}{2} \right) = \frac{e^{-\lambda}}{2} \lambda \frac{\pi}{2} \right\r
	$\pi(\lambda) = e^{-\beta\lambda} \cdot \lambda^{\infty-1} \cdot \beta^{\infty}$, $\lambda > 0$
	Toc
	therefore the posterior distribution is:
	$\pi(\lambda D) = \pi(\lambda y_1, y_2, \dots y_n) = f(y_1, y_2, \dots y_n \lambda) \pi(\lambda)$
	(6/27/1/11)
	meaning that I f (y, ye, yw) TT (N dx is independent of x
	and therefore is constant
	and therefore is constant $= e^{-N\lambda} \cdot \lambda \cdot \tilde{\xi}_{1} \cdot y_{1} \cdot e^{-B\lambda} \cdot \lambda^{\alpha-1} B^{\alpha} = \alpha e^{-\lambda(B+H)} \cdot \lambda^{(\frac{N}{2} y_{1}+\alpha-1)}$ $(\tilde{\chi}_{1}^{N} \cdot y_{1}) \qquad [\alpha]$
	AlD ngamma (ocp, Bp) where ocp = , Zy + oc- 1 and Bp = B+N
	therefore gamma distribution is a conjugate prior to the poisson distribution
	porssorialsmounter
	a new adults' republic and affic
	The same of the sa
	Land to the second to the seco
	the state of the s
	The second secon
	THE RESERVE AND ASSESSED FOR A STATE OF THE PARTY OF THE

Q4:	
97.	3000 = 1967
	Ordinal Encoding is when each category within a
	caregorical variable is assigned an integer value
	one- Hot Encoding is when each category within a
	categorical variable is separated and assigned
	binary values (O or I).
	Assertage marking 15-7/2 - Supplied to the second responsible for
	Utilizing ordinal encoding would not work (or atleast
	be incredibly inefficient) to utilize when it comes to
	Euclidean distance in kNN, when calculating the
	distance, you need to utilize numerical values that
	carry accurate weight. In terms of calculating
	Euclidean distance in KNN, the one-hot encoding
	approach would be optimal. This is because their
	numerical values actually carry weight when using
	them for calculations.
4	
-	

Q5:	Sum = 1967	
	count = 8000	
	total % with _ 1967 = 24.59 %	
	income >50k 8000 - 24.5476	
	Approximately 25% of individuals represented in the	
	training data have incomes greater than \$50k. A model	
	that achieves 70% accuracy is considered to be a good	1
	model based on online sources. Each data point has 12	
	dimensions excluding the increased dimensionality that	
	one-hot encodings bring. When including the increased	
	point has 85 dimensions (ignoring the id attribute and	
	closs label)	
)		
)		1
3		

	Debriefing
1.	I spent approximately 22 hours on this assignment, this was mainly because I am very unexperienced in Python, and this algorithm took a really long time for me to understand thoroughly.
2.	I would rate this assignment very difficult. The factors I mentioned above are the main reasoning for my rating. I also struggled slightly understanding how kaggle worked as well as interpretting what of was asking for. I am sure that students more familiar with Python & algorithms would rank this assignment lower difficulty than I did.
3.	I worked on this assignment alone, but I did view YouTube videos as well as our textbook chapters heavily in order to understand the kNN algorithm and k-fold cross validation.
4.	90%
5.	Not sure if this is an appropriate question to ask, but in comparison to this assignment, how much more difficult will the upcoming assignments be? (asking solely so that I can allot time accordingly, and ensure that I seek help well in advance)
40	