

# ReportMiner OCR Best Practices

FYdcfH jbyfjCdhW 7\UFUMYfFYVt[ b]h]cb fC 7F E W d U V ] ] h ] Y g U ` c k ` i g Y f g h c ` Y I h F U M X U H U Z f c a ` U b m i m d Y ` c Z g W b b Y X ` X c W a Y b h ` : c f V Y g h i f Y g i ` h g z h Y f Y V t [ b ] h ] c b ` Y b [ ] b Y ` b Y Y X g h c ` V Y h f U ] b Y X ` U b X ` h Y ` g Y h ] b [ g U X ` g h Y X " I g Y f g a i g h U g c ` h U \_ Y ` U Z Y k ` d f Y W i h ] c b g k ` \ Y b ` g W b b ] b [ ` h c ` Y b g i f Y Y U W ` \ X U F U M Y f ] g f Y V t [ b ] n Y X ` V t f f Y V t m i U b X ` a ] b ] a U z ` Z U b n z Y X ] h ] b [ ` ] g b Y Y X Y X ` d c g h i g W b "

K Y i j Y ` ] g h Y X ` U Z Y k ` ] a d c f h U b h V t b g ] X Y f U h ] c b g h c ` Y b g i f Y ` c d h ] a U ` C 7 F .

### Image and Font Size

- 8cW a Y b h g ` g \ c i ` X ` V Y ` g W b b Y X ` U h a ] b ] a i a ` \$ \$ ` X d ] f Y g c ` i h ] c b ` Z c f ` % \$ d h i h I h g j n Y z c f [ f Y U h Y f "
- : c f ` - d h Z c b h g j n Y ` c f ` g a U ` Y f z ( \$ \$ ! \* \$ \$ ` X d ] ] g f Y V t a a Y b X Y X "
- G U a Y j Y f h ] W ` U b X ` \ c f ] n c b h U ` f Y g c ` i h ] c b g ` g \ c i ` X ` ] X Y U ` m i V Y ` a U ] b h U ] b Y X "
- H n d Y ` c Z g W ] d h ] g U g c ` U Z W t c f k ` \ Y b ` X Y h Y f a ] b ] b [ ` ] X Y U ` h Y I h g j n Y ` V Y W i g y ` c Z X ] Z Y f Y b W ` ] b ` \ X U F U M Y f ` g j n Y .
  - G ] a d ` Y ` g W ] d h ] ` % V n t Y ` f b [ ` ] g \ z F i g g ] U b z 5 f U V ] W Y h W t F Y V t a a Y b X Y X ` h I h g j n Y ` 1 ` % d h
  - 7 c a d ` Y I ` g W ] d h ] ` & V n t Y g f T \ ] b Y g Y z > U d U b Y g Y z Y h W t F Y V t a a Y b X Y X ` h I h g j n Y ` 1 ` % d h A ] b ] a i a ` h I h g j n Y ` 1 ` % d h

### Color Recommendations

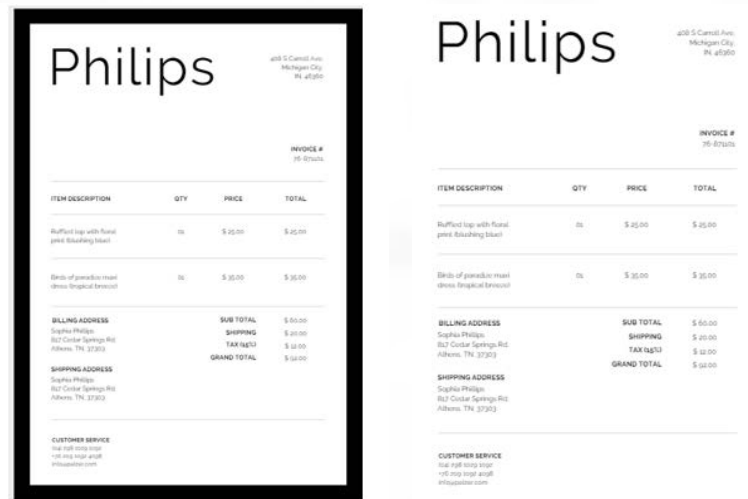
8cW a Y b h g k ] h ` V t a d ` Y I ` ` U n c i h g h U h V t b h U ] b ` V t ` c f Y X ` ] a U [ Y g c f ` V U W [ f c i b X g ` g \ c i ` X ` V Y ` V t b j Y f h Y X ` h c ` [ f U n g W Y ` Z c f V Y g h i f Y g i ` h g ` ; f U n g W Y ` ] a U [ Y g f Y j Y U ` a c f Y ` ] b Z c f a U h ] c b z c d h ] a ] n ] b [ ` C 7 F " " " " " "

## High-Quality Source Images

Ensure that the source documents are not damaged, wrinkled or discolored. For better results, use only high-quality. Lower quality images may require pre-processing techniques like binarization, contrast adjustments, and de-skewing.

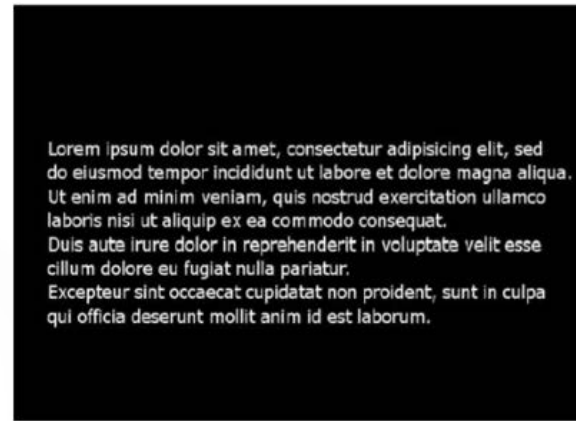
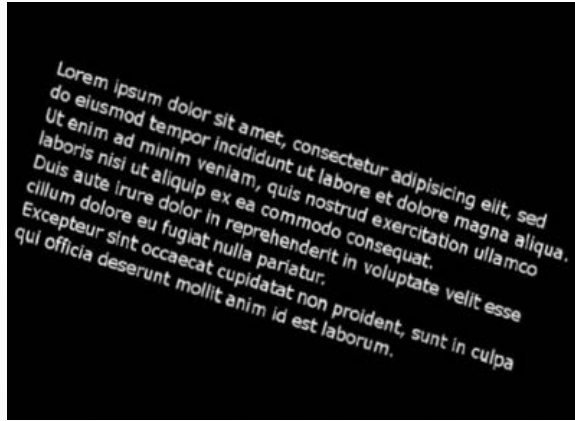
## Border Removal

Dark borders in source documents can be detected as extra characters by the OCR engine. We recommend removing them using OCR.



## Rotation/De-Skewing

The quality of OCR is reduced significantly if an image or page is skewed. De-skew or rotate the image so that is horizontally aligned for accurate OCR.



## Other Considerations

- The largest font size that ReportMiner's OCR can handle is 140pt.
- All image sizes can be processed.



k k k "UghYU'Vz'a  
7cbhUMi g'Zc'f'a cfY' ]bZcfa Uh]cb'cf'hc' fYei YghU'ZfYY' hf]U  
gUYg4 UghYU'Vz'a '""', , , !++! 5GH9F5

7cdmf][ \h&\$% '5ghYU'GcZk UFY' bVz'fdcfUHX' ]b'h\Y'I b]hYX'GHUhg'UbX'#cf'ch\Yf'Vz'i bhf]Yg"Ch\Yf'a Uf\_g'UFY'  
h\Y'dfcdYfmcZh\Y]f'fYgdYVW]j Y'ck bYfg"